

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 31, 2005, 01:52:59 ; Search time 174 Seconds  
(without alignments)  
555.690 Million cell updates/sec

Title: US-10-771-708-11  
Perfect score: 1354  
Sequence: 1 EQLKLVGFARFVSVADSVHR.....PPETGQNFVCKSVMYYYNE 250

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 2105692 seqs, 386760381 residues

Total number of hits satisfying chosen parameters: 2105692

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : A Geneseq\_16Dec04:\*  
1: geneseqp1980a:\*  
2: geneseqp1990a:\*  
3: geneseqp2000a:\*  
4: geneseqp2001a:\*  
5: geneseqp2002a:\*  
6: geneseqp2003a:\*  
7: geneseqp2003bs:\*  
8: geneseqp2004a:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	138	10.2	1638	4	ABB67889	AbB67889 Drosophila
2	125	9.2	758	4	ABB58337	AbB58337 Drosophila
3	121	8.9	744	4	ABB58426	AbB58426 Drosophila
4	120.5	8.9	692	4	ABB70535	AbB70535 Drosophila
5	116	8.6	790	4	AA667230	AA667230 Amino acid
6	113	8.3	833	4	ABB58185	AbB58185 Drosophila
7	105	7.8	812	3	AA550872	AA550872 Bovine pl
8	105	7.8	812	3	AA599595	AA599595 Bovine pl
9	105	7.8	812	4	AA667224	AA667224 Amino acid
10	104	7.7	810	4	AA667228	AA667228 Amino acid
11	102	7.5	810	2	AAW31169	AAW31169 Plasminog
12	101.5	7.5	790	2	AA660519	AA660519 Human 'G1
13	101	7.5	480	7	ABU64288	ABU64288 Human ang
14	101	7.5	563	5	ABB75942	ABB75942 Endotheli
15	101	7.5	566	2	AA702100	AA702100 A multifu
16	101	7.5	571	5	ABB75944	ABB75944 Angiotens
17	101	7.5	576	5	ABB75943	ABB75943 Angiotens
18	101	7.5	654	8	ABW83810	ABW83810 Human dia
19	101	7.5	672	8	ABW83809	ABW83809 Human dia
20	101	7.5	791	2	AAW34285	AAW34285 Human pla
21	101	7.5	791	3	AA550867	AA550867 Human pla
22	101	7.5	791	3	AA501887	AA501887 Human pla
23	101	7.5	791	3	AA599589	AA599589 Human pla
24	101	7.5	791	4	AA667223	AA667223 Amino acid
25	101	7.5	791	5	AAE23660	AAE23660 Human pla

26	101	7.5	791	6	ABG76087	ABG76087 Human pla
27	101	7.5	791	6	ADA08448	ADA08448 Mammalian
28	101	7.5	791	7	ABG75024	ABG75024 Glu-Plasmin
29	101	7.5	791	8	ADL97182	ADL97182 Plasminog
30	101	7.5	791	8	ADK23654	ADK23654 Human pla
31	101	7.5	791	8	ADQ14615	ADQ14615 Human pla
32	101	7.5	791	8	ADP67420	ADP67420 Human pla
33	101	7.5	791	8	ADQ28279	ADQ28279 Human pla
34	101	7.5	791	8	ADS20380	ADS20380 Human pla
35	101	7.5	807	2	AA820013	AA820013 PA mutant
36	101	7.5	810	2	AA808065	AA808065 Human pla
37	101	7.5	810	2	AA813221	AA813221 Human pla
38	101	7.5	810	2	AA813219	AA813219 R561G hum
39	101	7.5	810	2	AA812406	AA812406 R561G hum
40	101	7.5	810	2	AA813220	AA813220 R561S hum
41	101	7.5	810	2	AA812938	AA812938 Plasminog
42	101	7.5	810	2	AA834428	AA834428 Sequence
43	101	7.5	810	2	AA802114	AA802114 SEQ ID 77
44	101	7.5	810	2	AA808685	AA808685 Human pla
45	101	7.5	810	3	AA826690	AA826690 Human pla

ALIGNMENTS

RESULT 1  
ABB67889  
ID ABB67889 standard; protein, 1638 AA.  
XX  
AC ABB67889;  
XX  
DT 26-MAR-2002 (first entry)  
XX  
DE Drosophila melanogaster polypeptide SEQ ID NO 30459.  
XX  
KW Drosophila; developmental biology; cell signalling; insecticide;  
KM pharmaceutical.  
XX  
OS Drosophila melanogaster.  
XX  
PN WO200171042-A2.  
XX  
PD 27-SEP-2001.  
XX  
PF 23-MAR-2001; 2001WO-US009231.  
XX  
PR 23-MAR-2000; 2000US-0191637P.  
XX  
PR 11-JUL-2000; 2000US-00614150.  
XX  
PA (PEKE ) PE CORP NY.  
XX  
PI Venter JC, Adams M, Li PWD, Myers EW;  
XX  
DR WPI; 2001-656860/75.  
XX  
DR N-PSDB; ABL11992.  
XX  
PT New isolated nucleic acid detection reagent for detecting 1000 or more  
PT genes from Drosophila and for elucidating cell signaling and cell-cell  
PT interactions.  
XX  
PS Disclosure; SEQ ID NO 30459; 21pp + Sequence Listing; English.  
XX  
CC The invention relates to an isolated nucleic acid detection reagent  
CC capable of detecting 1000 or more genes from Drosophila. The invention is  
CC useful in developmental biology and in elucidating cell signalling and  
CC cell-cell interactions in higher eukaryotes for the development of  
CC insecticides, therapeutics and pharmaceutical drugs. The invention  
CC discloses genomic DNA sequences (AB146176-AB130511), expressed DNA  
CC sequences (AB101840-AB116175) and the encoded proteins (AB857737-  
CC AB872072). The sequence data for this patent did not form part of the  
CC printed specification, but was obtained in electronic format directly  
CC from WIPO at ftp.wipo.int/pub/published\_pcc\_sequences

SQ Sequence 1638 AA;

10.2%; Score 138; DB 4; Length 1638;

Query Match  
Best Local Similarity 21.2%; Pred. No. 0.00033;  
Matches 58; Conservative 43; Mismatches 100; Indels 72; Gaps 13;

QY 3 KILVGFARVVSADSVHRCISACIMAFDTFGFECESVMYYPVDAECILNTEDRLDRLDPDLF 62  
 Db 211 KLPISDIKEITLTAANR--SDGEDKCLN---EFSFVCRSANFPSTMRSCITLSFTTRTHPELM 266  
 QY 63 VDHEDTVITLDNNKACGCECHMFHFNFKISGILNDQFAIAQCYAPVTVYVAVGRQL 122  
 Db 267 EDDPNSD--YLENTCLNAE-----RRCDG--LAVFKEENKRL 300  
 QY 123 SD--BLDHSFEGELSECELCCTQRLSVTANDFNCKSFMYSNLTRGCVLSDEHS-RPLGR 179  
 Db 301 GGFPEVD-ITNNNTLBECCQTMG-----LRAEKYFCRSVERPDOSKOCITSEEDSISQKDD 354  
 QY 180 ANLAEPVGMWTFE-----SRGVPSFTRVPQMLVIG-FASE 213  
 Db 355 ISISSPTHHFYDLVCLDNGRANDYPDNSVTSHLFSGRRPDPAFQRYNRSRLGGEFHS 414  
 QY 214 VMENVPSTWCLDQCTSPPEPTGQNFVCKSVMY 246  
 Db 415 ITGR--SLSECLDECLRQ-----TSFQCRSAVY 440

## RESULT 2

ABB58337  
 ID ABB58337 standard; protein; 758 AA.

AC ABB58337;

DT 26-MAR-2002 (first entry)

XX Drosophila melanogaster polypeptide SEQ ID NO 1803.

XX Drosophila; developmental biology; cell signalling; insecticide;

KM pharmaceutical.

XX Drosophila melanogaster.

XX WO200171042-A2.

XX PD 27-SEP-2001.

PF 23-MAR-2001; 2001WO-US009231.

PR 23-MAR-2000; 2000US-0191637P.

PR 11-JUL-2000; 2000US-00614150.

XX (PEKE ) PE CORP NY.

PA Venter JC, Adams M, Li PWD, Myers EW;

XX MPI; 2001-656860/75.

XX N-PSDB; ABL02440.

PT New isolated nucleic acid detection reagent for detecting 1000 or more  
 PT genes from Drosophila and for elucidating cell signalling and cell-cell  
 PT interactions.

PS Disclosure; SEQ ID NO 1803; 21pp + Sequence Listing; English.

XX The invention relates to an isolated nucleic acid detection reagent  
 CC capable of detecting 1000 or more genes from Drosophila. The invention is  
 CC useful in developmental biology and in elucidating cell signalling and  
 CC cell-cell interactions in higher eukaryotes for the development of  
 CC insecticides, therapeutics and pharmaceutical drugs. The invention  
 CC discloses genomic DNA sequences (AB16176-AB130511), expressed DNA  
 CC sequences (AB101840-AB16175) and the encoded proteins (ABB57737-  
 CC ABB72072). The sequence data for this patent did not form part of the  
 CC printed specification, but was obtained in electronic format directly

CC from WIPO at ftp.wipo.int/pub/published\_pct\_sequences

XX SQ Sequence 758 AA;

9.2%; Score 125; DB 4; Length 758;

Query Match  
 Best Local Similarity 25.6%; Pred. No. 0.0025;  
 Matches 62; Conservative 25; Mismatches 93; Indels 62; Gaps 13;

QY 21 CTSACLNAPDVTGFECESVMYYPVDAECL--NTEDRLDRPLFPDEHEDTVITLDNNKA 78  
 Db 160 CHEAC-----QANESCSAVNY--ETGLCVMEFSTADQLP-GSLSSQYPVFTVYQKSCF 211  
 QY 79 G---CECHWFHFNFKTSGLINDQFAIAQCYAPVTVYVAVGRQLSPDLHSFEGLEL 135  
 Db 212 GVRPCKAMCITR-----VQGYRLPERAKASQSVATR 243  
 QY 136 SECELCCTQRLSVTANDFNCKSFMYSNLTRGCVLSD-EKSRPLGRANLAEPVGMWTFESR 194  
 Db 244 RDCIEIC-----LGETEFCSRANYAHSGLCELSDMRITLSDANIAAYDGA DYLENN 298  
 QY 195 GV--PS-----FTRVPQMLLVGFASFVMEVNSVTWCLDQCTSPPEPTGQNFVCKSVMY 248  
 Db 299 CAEPFSKLCFRRVAGRIKTVDS-VHONVQTLDECRDLCTAP-----FRCHS--YDY 349  
 QY 249 NE 250  
 Db 350 NE 351

## RESULT 3

ABB58426  
 ID ABB58426 standard; protein; 744 AA.

AC ABB58426;

DT 26-MAR-2002 (first entry)

XX Drosophila melanogaster polypeptide SEQ ID NO 2070.

XX Drosophila; developmental biology; cell signalling; insecticide;

KM pharmaceutical.

XX Drosophila melanogaster.

XX WO200171042-A2.

XX PD 27-SEP-2001.

PF 23-MAR-2001; 2001WO-US009231.

PR 23-MAR-2000; 2000US-0191637P.

PR 11-JUL-2000; 2000US-00614150.

XX (PEKE ) PE CORP NY.

PA Venter JC, Adams M, Li PWD, Myers EW;

XX MPI; 2001-656860/75.

XX N-PSDB; ABL02529.

PT New isolated nucleic acid detection reagent for detecting 1000 or more  
 PT genes from Drosophila and for elucidating cell signalling and cell-cell  
 PT interactions.

PS Disclosure; SEQ ID NO 2070; 21pp + Sequence Listing; English.

XX The invention relates to an isolated nucleic acid detection reagent  
 CC capable of detecting 1000 or more genes from Drosophila. The invention is  
 CC useful in developmental biology and in elucidating cell signalling and  
 CC cell-cell interactions in higher eukaryotes for the development of  
 CC insecticides, therapeutics and pharmaceutical drugs. The invention  
 CC discloses genomic DNA sequences (AB16176-AB130511), expressed DNA  
 CC sequences (AB101840-AB16175) and the encoded proteins (ABB57737-

CC ABB72072). The sequence data for this patent did not form part of the  
CC printed specification, but was obtained in electronic format directly  
CC from WIPO at ftp.wipo.int/pub/published\_pct\_sequences  
XX  
SQ Sequence 744 AA;

Query Match 8.9%; Score 121; DB 4; Length 744;

Best Local Similarity 23.4%; Pred. No. 0.0061;  
Matches 64; Conservative 29; Mismatches 102; Indels 78; Gaps 14;

```
QY 5 LVGFAREVNSA-----DS-----VHRCISACINAEPTGFECESVMYYPVDAEC 48
D 117 LVGF--EIVTGVSFAPBKLMDSPGTMLMTDCLDTCKR-----NKTGOSVNY--ETGLC 167
QY 49 ILNTEDRIDRPD-LFVDEHEDTVITYLDNNKAG---CECHWHFDPNKTGILNDQFALIA 104
D 168 VLFSAHADQLPGALTKSQFPVFTIYAQRKCLAVKCSBAMWYDR----- 211
QY 105 QCVAPVYQYVAVEGRQSLDELHSEFGLSECEELCTQRLSVTANPFNCKSPFVSNLT 164
D 212 -----VQNYKLKTEYKRTVSVASRRECELC-----LGENDTGRSANYDRTS 254
QY 165 RSCVLSDESRSPPLGRANLAEV-PGWTYFESRGVP-----SFTRVPMILVGFASFVME 217
D 255 GACBELSELDRILTAGSQAFQVNDGSDYLENHCVDSPNKLCEFKRLPRLIKTVDS-VYQE 313
QY 218 VPSVTMCLDQCTSPPEPTGQNFVCKSVMYRYNE 250
D 314 VSSIDECRELCLNSP-----YRCHS--YDYND 338
```

## RESULT 4

ABB70535  
ID ABB70535 standard; protein; 692 AA.

XX AC ABB70535;

XX DT 26-MAR-2002 (first entry)

DE Drosophila melanogaster polypeptide SEQ ID NO 38397.

XX Drosophila; developmental biology; cell signalling; insecticide;

KW pharmaceutical.

XX Drosophila melanogaster.

PN WO200171042-A2.

PD 27-SEP-2001.

XX 23-MAR-2001; 2001WO-US009231.

PR 23-MAR-2000; 2000US-0191637P.

PR 11-JUL-2000; 2000US-00614150.

XX (PEKE ) PE CORP NY.

PI Venter JC, Adams M, Li PWD, Myers EW;

DR WPI; 2001-656860/75.

DR N-PSDB; ABL14638.

PT New isolated nucleic acid detection reagent for detecting 1000 or more  
PT genes from Drosophila and for elucidating cell signalling and cell-cell  
PT interactions.

PS Disclosure; SEQ ID NO 38397; 21pp + Sequence Listing; English.

CC The invention relates to an isolated nucleic acid detection reagent  
CC capable of detecting 1000 or more genes from Drosophila. The invention is  
CC useful in developmental biology and in elucidating cell signalling and  
CC cell-cell interactions in higher eukaryotes for the development of  
CC insecticides, therapeutics and pharmaceutical drugs. The invention

CC discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA  
CC sequences (ABL01840-ABL16175) and the encoded proteins (AB57737-  
CC ABB72072). The sequence data for this patent did not form part of the  
CC printed specification, but was obtained in electronic format directly  
CC from WIPO at ftp.wipo.int/pub/published\_pct\_sequences  
XX  
SQ Sequence 692 AA;

Query Match 8.9%; Score 120.5; DB 4; Length 692;

Best Local Similarity 19.9%; Pred. No. 0.0062;  
Matches 53; Conservative 35; Mismatches 92; Indels 87; Gaps 9;

```
QY 3 KILVFAREVNSADSVHRCISACINAEPTGFECESVMYYPVDAECILNTEDRIDRPLF 62
D 118 KVIKGLDNALITYSTKACISACIN---ERRFVRSVEYDNNKCVLSDDRSSGQFV 174
QY 63 VDEHEDTVITYDNNC---ACCECHWHFDPNKTGILNDQFALIAQCVAPVYQYVAE 118
D 175 QLVDAQGTDPYENLCLKPAQACKNNRSFGNSQKGVSEK-----VAQYGL- 221
QY 119 GRQSLDELHSEFGLSECEELCTQRLSVTANPFNCKSPFVSNLT 165
D 222 -----HYTDKEQLVTSSESACRLACEISEFCLCRFLYLQPGSQYNCRLLYLDH 272
QY 166 -----SCVLSDBR-----SRPLGRANLAEVPGWTYFESR-----GVPSFTRV 204
D 273 KTLPDGSESTYLNHRPRLIDHEPFIQ-----YENQCEKRAAGLAGASPTOT----- 318
QY 205 MLVGFASFVMEANVPSVTMCLDQCTSP 231
D 319 -----LDKIDTLPLVSLDTIEDP 335
```

## RESULT 5

AAG67230  
ID AAG67230 standard; protein; 790 AA.

XX AC AAG67230;

XX DT 13-NOV-2001 (first entry)

DE Amino acid sequence of pig plasminogen.

XX

KW Angiostatin; plasminogen; sulphydryl donor; angiogenesis; tumour;

KW angiogenic disease; neoplastic disease; connective tissue disorder;

KW rheumatoid arthritis; atherosclerosis; ocular angiogenic disease;

KW diabetic retinopathy; corneal graft rejection; cardiovascular disease;

KW cerebral vascular disease; diabetes; immune disorder;

KW chronic inflammation; autoimmunity.

XX

OS Sus scrofa.

XX WO200158921-A2.

XX 16-AUG-2001.

PR 08-FEB-2001; 2001WO-US004021.

PR 08-FEB-2000; 2000US-00500397.

XX (NOUN ) UNIV NORTHWESTERN.

PI Soft G, Gately ST, Twardowski P;

DR WPI; 2001-550019/61.

PT Producing angiostatin for treating angiogenic diseases involves  
PT contacting plasminogen with plasminogen activator and sulphydryl donor  
PT simultaneously, or producing plasmin which is contacted with sulphydryl  
PT donor.  
PS Disclosure; Page 77-80; 101pp; English.





PF 08-FEB-2001; 2001W0-US004021.  
 XX  
 XX 08-FEB-2000; 2000US-00500397.  
 XX  
 XX (NOUN ) UNIV NORTHWESTERN.  
 PA  
 PI Soff G, Gately ST, Twardowski P;  
 XX  
 XX WPI; 2001-550019/61.  
 DR  
 XX  
 PT Producing angiostatin for treating angiogenic diseases involves  
 PT contacting plasminogen with plasminogen activator and sulfhydryl donor  
 PT simultaneously, or producing plasmin which is contacted with sulfhydryl  
 PT donor.  
 PS Disclosure; Page 80-84; 101pp; English.  
 XX  
 XX The specification describes a method for generating angiostatin in vitro.  
 CC The method comprises contacting plasminogen with a sulfhydryl donor, or  
 CC culturing cells capable of producing plasminogen activator in conditioned  
 CC culture medium (CCM) and contacting the CCM with plasminogen. Angiostatin  
 CC produced by method of the invention is useful for treating animals with  
 CC angiogenesis diseases. It is useful for treating an angiogenic disease  
 CC such as neoplastic diseases (e.g. tumours and tumour metastasis), benign  
 CC tumours (e.g. hemangiomas, acoustic neuromas, etc), connective tissue  
 CC disorders (e.g. rheumatoid arthritis and atherosclerosis), ocular  
 CC angiogenic diseases (e.g. diabetic retinopathy, corneal graft rejection,  
 CC etc), cardiovascular diseases, cerebral vascular diseases, diabetes-  
 CC associated diseases and immune disorders (e.g. chronic inflammation and  
 CC autoimmunity). The present sequence represents a plasminogen  
 CC  
 XX  
 SQ Sequence 812 AA;  
 Query Match 7.8%; Score 105; DB 4; Length 812;  
 Best Local Similarity 18.3%; Pred. No. 0.28; Indels 120; Gaps 17;  
 Matches 55; Conservative 42; Mismatches 83;  
 QY 5 LVGPAREVVSADSVHRCLSACLNAPDTFGFCESEVMYYPVDAECILNTEDRDLRPLFVD 64  
 DB 40 LLSLSRKNLAGRSVEDCAAKCEETD--FVCRAPFYHSKEQCVMMAENSKVTP--VF 93  
 QY 65 EHEDTVIYLDN-----NC-----AGECH-W-----HFDNFKSGILNQ 98  
 DB 94 RMRDVLVLEKRYLLRECKTGNGQYRTGTAETKSGYTCQKMSYSPHPK-----SE 147  
 QY 99 QPAAIA-----QCYAP-----YVTOYVAVGROLSDLDHSFEGLEISECEELCTO- 144  
 DB 148 KFPPLAGEENYCRNPNDENGFWCYTT-----DPDKRYVYCCIPECEDCMHC 195  
 QY 145 -----RLSVTANDFNCKS-----FMYSNLTRS-CVLSDESRPLGRAN 181  
 DB 196 SGENYEGKIAKTAKTSGRDQAMDSQSPHAGYIPSKFPNNKLNKMYCRNPDGRRP----- 250  
 QY 182 LAEPVMTYFESRGVSPFRVPQMLLVGFASFVMEVNPSTVMKLDQCTSPPEPTQNFVC 241  
 DB 251 -----W-----CFTTDPQ-----KRWFCIDP-----RCTTPPSSGPTYQC 282  
 RESULT 10  
 AAG67228  
 ID AAG67228 standard; protein; 810 AA.  
 AC AAG67228;  
 XX  
 XX 13-NOV-2001 (first entry)  
 DE Amino acid sequence of monkey plasminogen.  
 XX  
 XX Angiostatin; plasminogen; sulfhydryl donor; angiogenesis; tumour;  
 KW angiogenic disease; neoplastic disease; connective tissue disorder;  
 KW rheumatoid arthritis; atherosclerosis; ocular angiogenic disease;  
 KW diabetic retinopathy; corneal graft rejection; cardiovascular disease;  
 KW cerebral vascular disease; diabetes; immune disorder;

KW chronic inflammation; autoimmunity.  
 XX  
 XX Macaca mulatta.  
 OS  
 XX WO200158921-A2.  
 PN  
 XX 16-AUG-2001.  
 PD  
 XX  
 XX 08-FEB-2001; 2001W0-US004021.  
 PF  
 XX  
 PR 08-FEB-2000; 2000US-00500397.  
 XX  
 XX (NOUN ) UNIV NORTHWESTERN.  
 PA  
 PI Soff G, Gately ST, Twardowski P;  
 XX  
 XX WPI; 2001-550019/61.  
 DR  
 XX  
 XX The specification describes a method for generating angiostatin in vitro.  
 CC The method comprises contacting plasminogen with a sulfhydryl donor, or  
 CC culturing cells capable of producing plasminogen activator in conditioned  
 CC culture medium (CCM) and contacting the CCM with plasminogen. Angiostatin  
 CC produced by method of the invention is useful for treating animals with  
 CC angiogenesis diseases. It is useful for treating an angiogenic disease  
 CC such as neoplastic diseases (e.g. tumours and tumour metastasis), benign  
 CC tumours (e.g. hemangiomas, acoustic neuromas, etc), connective tissue  
 CC disorders (e.g. rheumatoid arthritis and atherosclerosis), ocular  
 CC angiogenic diseases (e.g. diabetic retinopathy, corneal graft rejection,  
 CC etc), cardiovascular diseases, cerebral vascular diseases, diabetes-  
 CC associated diseases and immune disorders (e.g. chronic inflammation and  
 CC autoimmunity). The present sequence represents a plasminogen  
 CC  
 XX  
 SQ Sequence 810 AA;  
 Query Match 7.7%; Score 104; DB 4; Length 810;  
 Best Local Similarity 18.1%; Pred. No. 0.36; Indels 118; Gaps 14;  
 Matches 54; Conservative 38; Mismatches 89;  
 QY 5 LVGPAREVVSADSVHRCLSACLNAPDTFGFCESEVMYYPVDAECILNTEDRDLRPLFVD 64  
 DB 33 LFSITKQDLAGSISECAKCEBER--FTCRSQHNSKEQCVMMAE--NRKSSIVF 86  
 QY 65 EHEDTVIYLDNNCAGCECHMFDFNFKTS-----GILNDQFAIAAGCY--APVTOYV 115  
 DB 87 RMRDVLVLP-----EKVYLSECKTGNGKMYRGTMKTRGILTCQKMSSTSPH----- 133  
 QY 116 AVEGQLSDELDSHSEGLE-----LSSCEELCTO- 144  
 DB 134 ---RPTSPATHBSEGLEENYCRNPNDQGPWCYTTDPERPDDYCDIPCECECHMCS 189  
 QY 145 -----RLSVTANDFNCKS-----FMYSNLTRS-CVLSDESRPLGRANL 182  
 DB 190 GENYDGKISKTMISGLECCAMDSQSPHAGYIPSKFPNNKLNKMYCRNPDGRRP----- 243  
 QY 183 AEVPGMTYFESRGVSPFRVPQMLLVGFASFVMEVNPSTVMKLDQCTSPPEPTQNFVC 241  
 DB 244 -----W-----CFTTDPN-----KRWELCIDP-----RCTTPPSSGPTYQC 275  
 RESULT 11  
 AAW31169  
 ID AAW31169 standard; protein; 810 AA.  
 AC AAW31169;  
 XX  
 XX 26-FEB-1998 (first entry)

XX Plasmidogen protein for production of recombinant plasmidogen.  
 DE Plasmidogen, alpha-1-antitrypsin; AAT; argaspirin; recombinant;  
 XX stabilising protein.  
 KW Homo sapiens.  
 XX  
 OS  
 PH Key Location/Qualifiers  
 FT Misc-difference 60 /note= "encoded by ACC"  
 FT Misc-difference 191 /note= "encoded by GAA"  
 FT Misc-difference 223 /note= "encoded by AAA"  
 FT Misc-difference 280 /note= "encoded by GCT"  
 FT Misc-difference 291 /note= "encoded by GSA"  
 FT Misc-difference 395 /note= "encoded by GSA"  
 FT Misc-difference 503 /note= "encoded by CAG"  
 FT Misc-difference 607 /note= "encoded by TCT"  
 FT Misc-difference 615 /note= "encoded by CAG"  
 FT Misc-difference 658 /note= "encoded by TTG"  
 FT Misc-difference 709 /note= "encoded by GCT"  
 FT US5648254-A.  
 XX  
 PV 15-JUL-1997.  
 PD 14-JUL-1994; 94US-00275076.  
 XX  
 PP 15-JAN-1988; 88US-00144357.  
 PR 04-DEC-1989; 89US-00445302.  
 PR 28-OCT-1991; 91US-00785865.  
 XX  
 PA (ZYMO ) ZYMOGENETICS INC.  
 PI Kumar AA, Mulvihill ER;  
 DR WPI; 1997-372063/34.  
 DR N-PSDB; AAT89686.  
 XX  
 PT Production of recombinant plasmidogen - by co-expression with plasmidogen  
 PT -processing or -stabilising protein.  
 XX  
 PS Example 2; Fig 6A-D; 32pp; English.  
 XX  
 CC This is the protein plasmidogen. The encoding cDNA is used in a new  
 CC process for the production of plasmidogen where a first DNA sequence  
 CC encoding plasmidogen and at least one additional DNA sequence encoding a  
 CC protein that processes or stabilises the plasmidogen is introduced into a  
 CC eukaryotic host cell. The protein is selected from alpha-1-antitrypsin  
 CC (AAT) and its variants and Argaspirin. Both the DNA sequences are  
 CC operably linked to transcriptional promoter and terminator sequences. The  
 CC host cell is cultured under conditions that allow the DNA sequences to be  
 CC expressed and the recombinant plasmidogen is isolated from the host cell.  
 CC Co-expression of plasmidogen and the protein gives increased yields of  
 CC undegraded plasmidogen  
 CC  
 CC Sequence 810 AA;  
 SQ

Query Match 7.5%; Score 102; DB 2; Length 810;  
 Best Local Similarity 17.7%; Pred. No. 0.57; Indels 108; Gaps 12;  
 Matches 52; Conservative 45; Mismatches 89; Indels 108; Gaps 12;  
 5 LVGFAREVVSADSVARCLSLAIAFDTEGCECEVYVVDACILNTEDR-----LDRP 59

Db 33 LFSVTKKQLGAGSIEBCAKCEDEB---FYCRAFOYHSKEQOCVIMENRKSIIIMR 89  
 Qy 60 DLFDDEHEDTIVYLDNNAGCECHMHFNFETSGINDQQAIAQCAAPVTQVAVNEG 119  
 Db 90 DVLFEEKK---VYISECKTGKNYRGTMSTKNGITCKWSSST---SPH----- 133  
 Qy 120 RQLSDELDHSPEGLE-----LSECEELCTQ----- 144  
 Db 134 RPRSPATHPSBEGLEBNYCRPNDDPGQPCYTTDPREKRYCYCLIEBEEBCMHSGGNY 193  
 Qy 145 --RLSVTANDFNCKS-----FMYSNLTRSCVLSDBSRHPLGRANLAEPVG 187  
 Db 194 DGKISKTWGSGLECOAMDSPHAGYIPSCFPNKLKKNYCRNPER-----ELRP 243  
 Qy 188 WTEFESKGVSPFTVPQMLVGFASPVVENYPSVTMLCDQCTSPPTGQNFVC 241  
 Db 244 W-----CFTYDPN-----KAMELCIDP-----RCTTPPSSGPTYQC 275

RESULT 12

AA60519 standard; protein; 790 AA.  
 AA60519;  
 AC 25-MAR-2003 (revised)  
 DT 22-MAR-1995 (first entry)  
 XX Human 'Glu' plasmidogen.  
 XX  
 KW Serine protease; Factor-Xa; recognition site; plasmidogen; kringle;  
 KW fusion protein cleavage; protein folding; primer;  
 KW polymerase chain reaction; amplification.  
 XX  
 OS Homo sapiens.  
 XX  
 PN W09418227-A2.  
 PD 18-AUG-1994.  
 XX  
 PP 04-FEB-1994; 94WO-DK000054.  
 XX  
 PR 04-FEB-1993; 93DK-00000130.  
 PR 05-FEB-1993; 93DK-00000139.  
 PR 03-DEC-1993; 93WO-GB002492.  
 XX  
 PA (DENZ-) DENZYME APS.  
 XX  
 PI Thogersen HC, Hollet TL, Etzerodt M;  
 DR WPI; 1994-279681/34.  
 XX  
 PT Refolding of polypeptide molecules - using a cyclic process involving  
 PT denaturing and renaturing conditions to produce a correctly folded prod.  
 XX  
 PS Disclosure: Page 148-50; 202pp; English.  
 XX  
 CC cDNA encoding kringle domains 1 and 4 of human plasmidogen (full sequence  
 CC given in AA60519) was PCR amplified using primers given in AAQ71268-71.  
 CC Amplified cDNA was linked to a sequence encoding the Factor-Xa cleavage  
 CC site (given in AA60503), subcloned in vector pLICIMCH6 so that it was  
 CC linked to a hexahistidine-encoding sequence and expressed in E. coli  
 CC QY13. The fusion protein was purified on an Ni2+-activated NTA-agarose  
 CC column. A cyclic procedure was used to obtain correctly folded  
 CC recombinant protein. (Updated on 25-MAR-2003 to correct PN field.)  
 CC  
 CC Sequence 790 AA;  
 SQ

Query Match 7.5%; Score 101.5; DB 2; Length 790;  
 Best Local Similarity 18.0%; Pred. No. 0.62;  
 Matches 53; Conservative 42; Mismatches 90; Indels 109; Gaps 13;

```

QY 5 LVGFAREVVSADSVHRCISACLNAPDTFGFCECEVVMYPVDAECILNTEDR----LDRPD 60
DB 14 LFSVTKKQJGAGSIEECAKCEDEE---FTGRAFOYHSKEQOCYVMAENKRSIITMRD 70
QY 61 LFDVDEHEDTVIYLDNNAGCECHMHPDNFKTSGLINDQCPAIAAOCYAPVYTVQYAVEGR 120
DB 71 VVLFPEKK---VYLSECKTGNKXVGTGSKTNGITQCKMSST---SPH-----R 114
QY 121 QLSDELDSFEGLF-----LSRCEELCTQ----- 144
DB 115 PRFSPATHPEGLSEENYCRNPNDPQGPWCYTTDPEKRYDYCDILECEEECHMGSGENYD 174
QY 145 -RLSVTANDFNCKS-----FMYSNLTRS--CVLSDESRRLPGRANLAEPVG 187
DB 175 GKISKTMSGLECOAMDQSOPHAGYIPSKFPNNKLKXVCRNPDRELRP----- 223
QY 188 WTYFESRGVPSFTFVPQMLVGFASFVMEVNSVTMCLDOCTSPPEPTQGNFVC 241
DB 224 W-----CFTTDPN-----KRWELCDIP-----RCTTPPSSGPTYQC 255

```

## RESULT 13

ABU64288 standard; protein; 480 AA.

ABU64288;

11-MAR-2004 (first entry)

Human angiotensin protein.

Vector; rAAV; recombinant adeno-associated viral vector;

anti-angiogenesis; PEDF; angiogenesis; eye disorder; blindness;

retinal degeneration; macular degeneration; neovascularisation;

ophthalmological.

Homo sapiens.

WO2003080648-A2.

02-OCT-2003.

20-MAR-2003; 2003WO-US008667.

20-MAR-2002; 2002US-0366114P.

(UYFL ) UNIV FLORIDA RES FOUND INC.

(UYCO ) UNIV JOHNS HOPKINS.

Hauswirth WW, Campochiaro PA, Berns KI;

WPI; 2003-779243/73.

Novel adeno-associated viral vector comprising polynucleotide encoding pigment epithelium-derived factor, useful for treating choroidal neovascularization, blindness, loss of vision.

Claim 14; Page 37; opp; English.

The present invention relates to an adeno-associated viral (AAV) vector comprising a polynucleotide that comprises a nucleic acid segment that encodes a choroidal or ocular neovascularisation inhibitory polypeptide operably linked to a promoter that expresses the segment to produce the polypeptide in a selected mammalian host cell. Such a vector is useful for providing a choroidal or ocular neovascularisation inhibitory polypeptide to a mammal, for use in the therapy of ocular neovascularisation, choroidal neovascularisation, retinal neovascularisation, age-related macular degeneration, visual impairment, ocular dysfunction, loss of vision, retinopathy, or blindness in a human. The present sequence is a protein shown in the exemplification of the invention

Sequence 480 AA;

Query Match 7.5%; Score 101; DB 7; Length 480;  
Best Local Similarity 18.0%; Pred. No. 0.36;  
Matches 53; Conservative 42; Mismatches 90; Indels 110; Gaps 13;

```

QY 5 LVGFAREVVSADSVHRCISACLNAPDTFGFCECEVVMYPVDAECILNTEDR----LDRP 59
DB 33 LFSVTKKQJGAGSIEECAKCEDEE---FTGRAFOYHSKEQOCYVMAENKRSIITMR 89
QY 60 LFDVDEHEDTVIYLDNNAGCECHMHPDNFKTSGLINDQCPAIAAOCYAPVYTVQYAVEG 119
DB 90 DVVLFPEKK---VYLSECKTGNKXVGTGSKTNGITQCKMSST---SPH----- 133
QY 120 QLSDELDSFEGLF-----LSRCEELCTQ----- 144
DB 134 PRFSPATHPEGLSEENYCRNPNDPQGPWCYTTDPEKRYDYCDILECEEECHMGSGENY 193
QY 145 -RLSVTANDFNCKS-----FMYSNLTRS--CVLSDESRRLPGRANLAEPVG 186
DB 194 DGKISKTMSGLECOAMDQSOPHAGYIPSKFPNNKLKXVCRNPDRELRP----- 243
QY 187 WTYFESRGVPSFTFVPQMLVGFASFVMEVNSVTMCLDOCTSPPEPTQGNFVC 241
DB 244 W-----CFTTDPN-----KRWELCDIP-----RCTTPPSSGPTYQC 275

```

## RESULT 14

ABB75942 standard; protein; 563 AA.

ABB75942;

01-JUL-2002 (first entry)

Endothelial cell growth inhibitor, angiogenesis.

Plasminogen; angiogenesis; inhibitor; tumour; antitumour;

cytostatic; antiproliferative; dermatological; ophthalmological;

antidiabetic; antiarthritic; vulnerability; antitumor; antithrombotic;

contraceptive; gene therapy; human.

Homo sapiens.

Synthetic.

Key

Peptide

Peptide

Peptide

Peptide

Peptide

Peptide

Peptide

Peptide

Peptide

Peptide

Peptide

Peptide

Peptide

Peptide

Peptide

Peptide



CC The present sequence is the protein sequence of a novel recombinant  
CC endothelial cell growth inhibitor, termed angioquiescin, which comprises  
CC the signal peptide and pre-activation peptide of plasminogen (see  
CC ABB75939) in association with Kringle 1-5. Angioquiescin may also  
CC include a C-terminal tumour-targeting peptide (see ABB75943-44).  
CC Angioquiescin, nucleic acids encoding it, vectors and host cells are used  
CC in the preparation of a medicament for treating a tumour by preventing  
CC angiogenesis (claimed). In claimed methods, the peripheral and tumour-  
CC infiltrating lymphocytes, hepatocytes, epidermal cells, myocytes or other  
CC somatic cells of a patient in need of anti-angiogenic therapy are  
CC transfected with a viral (especially retroviral) vector comprising an  
CC angioquiescin nucleic acid, and the transfected cells are then expanded  
CC and implanted into the patient. Angioquiescin is useful for inhibiting  
CC cancer metastasis, and for treating angiogenesis associated conditions  
CC such as tumour growth e.g., cancer, diabetes, etc. It is also useful for  
CC treating haemangioma, solid tumours, leukaemia, psoriasis, scleroderma,  
CC arteriovenous malformations, rubecosis, neovascular glaucoma, diabetic  
CC retinopathy, arthritis, diabetic neovascularisation, macular  
CC degeneration, wound healing, peptic ulcer, intestinal adhesions,  
CC atherosclerosis, fractures, keloids, vasculogenesis, haematopoiesis or  
CC cat scratch fever. The protein can also be used as a birth control agent  
CC by preventing the vascularisation required for embryo implantation. The  
CC half maximal concentration (EC50) of angioquiescin for inhibiting  
CC endothelial cell proliferation is about 50 pM, compared with 100 nM for  
CC angioestatin  
XX  
SQ Sequence 563 AA;

Query Match 7.5%; Score 101; DB 5; Length 563;  
Best Local Similarity 18.0%; Pred. No. 0.44;  
Matches 53; Conservative 42; Mismatches 90; Indels 110; Gaps 13;

OY 5 LVGFARREVSADSVHRCISACLNAFDTFGFECESVMYYPVDAECLINTEDR-----LDRP 59  
DB 33 LFSVTKKQLGAGSIEBCAKCEDEE---FTCRAFQYHSKEQCQVIMENRKSIIIMR 89  
OY 60 DLFDVDEHDTYIYLDNNCAGECECHMFNFTSGILNQQFALIAQCYAPYTVQVAVEG 119  
DB 90 DVVLFEKK---VYLSECKTGKNGKRYGTMSKTKNGITCQKMSST---SPH----- 133  
OY 120 RQLSDELDSPEGLF-----LSECEBELCTQ----- 144  
DB 134 RPRFSRATHPSGGLSENYCRPNDDPGQPMCYTTDPEKRYDYCDILECEBECMHGSGENY 193  
OY 145 --RLSVTANDPNCKS-----FMYSNLTRS--CVLSDESRRLPYGRANLAEVP 186  
DB 194 DGKISKTSGLEGCQAMDSQSPHAGYIPSKFPPNKVLKKNYCRNPDRELRP----- 243  
OY 187 GWTYPESRGVPSFTFRVPMQLVGFASFVMEVNPVSTWKLDOCTSPRPETGQNFVC 241  
DB 244 -W-----CFTTDPN-----KRWELCDIP-----RCTTTPPSSGPTYOQ 275

RESULT 15  
AA02100  
ID AAY02100 standard: protein; 566 AA.

XX  
DT 16-JUL-1999 (first entry)

XX A multifunctional protein of the invention.

KM Angiostatin; endostatin; interferon; thrombospondin;  
KM interferon-inducible protein; platelet factor 4; anti-angiogenic;  
KM anti-tumor; multifunctional protein; angiogenic-mediated disease; cancer;  
KM diabetic retinopathy; macular degeneration; arthritis;  
KM tumor cell production.

XX Synthetic.  
OS Homo sapiens.  
XX  
XX  
PN WO9916889-A1.

XX  
PD 08-APR-1999.  
XX  
XX 30-SEP-1998; 98WO-US020464.  
XX  
XX 01-OCT-1997; 97US-0060609P.  
XX  
PA (SEAR ) SEARLE & CO G D.  
XX  
XX Bolanowski MA, Caparon MH, Casperon GF, Gregory SA, Klein BK;  
P1 McKeate JF;  
XX WPI; 1999-255098/21.  
XX  
PT New multifunctional proteins useful for treating angiogenic-mediated  
PT diseases.

XX Claim 5; Page 95-96; 121pp; English.

CC The specification describes multifunctional proteins which comprise  
CC combinations of angioestatin, endostatin, interferon, thrombospondin,  
CC interferon-inducible protein and platelet factor 4, and have anti-  
CC angiogenic and/or anti-tumor activity. The multifunctional protein may  
CC exhibit useful properties such as having similar or greater biological  
CC activity when compared to a single factor or by having improved half-life  
CC or decreased adverse side effects, or a combination of these properties.  
CC The proteins can be used for treating an angiogenic-mediated disease,  
CC e.g., cancer, diabetic retinopathy, macular degeneration, or arthritis.  
CC They can also be used for inhibiting the production of tumor cells  
CC (characteristic of lung, breast, ovarian, prostate, pancreatic, gastric,  
CC colon, renal, bladder cancers; melanoma, hepatoma, sarcoma and lymphoma)  
CC in a patient and for inhibiting tumor growth. The present sequence  
CC represents a multifunctional protein of the invention  
XX  
SQ Sequence 566 AA;

Query Match 7.5%; Score 101; DB 2; Length 566;  
Best Local Similarity 18.0%; Pred. No. 0.45;  
Matches 53; Conservative 42; Mismatches 90; Indels 110; Gaps 13;

OY 5 LVGFARREVSADSVHRCISACLNAFDTFGFECESVMYYPVDAECLINTEDR-----LDRP 59  
DB 33 LFSVTKKQLGAGSIEBCAKCEDEE---FTCRAFQYHSKEQCQVIMENRKSIIIMR 89  
OY 60 DLFDVDEHDTYIYLDNNCAGECECHMFNFTSGILNQQFALIAQCYAPYTVQVAVEG 119  
DB 90 DVVLFEKK---VYLSECKTGKNGKRYGTMSKTKNGITCQKMSST---SPH----- 133  
OY 120 RQLSDELDSPEGLF-----LSECEBELCTQ----- 144  
DB 134 RPRFSRATHPSGGLSENYCRPNDDPGQPMCYTTDPEKRYDYCDILECEBECMHGSGENY 193  
OY 145 --RLSVTANDPNCKS-----FMYSNLTRS--CVLSDESRRLPYGRANLAEVP 186  
DB 194 DGKISKTSGLEGCQAMDSQSPHAGYIPSKFPPNKVLKKNYCRNPDRELRP----- 243  
OY 187 GWTYPESRGVPSFTFRVPMQLVGFASFVMEVNPVSTWKLDOCTSPRPETGQNFVC 241  
DB 244 -W-----CFTTDPN-----KRWELCDIP-----RCTTTPPSSGPTYOQ 275

Search completed: March 31, 2005, 02:03:03  
Job time : 177 secs

This page blank (uspio)

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: March 31, 2005, 01:57:45 ; Search time 43 Seconds  
(without alignments)  
434.006 Million cell updates/sec

Title: US-10-771-708-11

Perfect score: 1354  
Sequence: 1 EQLIVGFARFVVSADSVHR.....PPETGQNFVCKSVMYYYNE 250

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 513545 seqs, 74649064 residues

Total number of hits satisfying chosen parameters: 513545

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Issued Patents AA: \*  
1: /cgn2\_6/ptodata/1/1aa/5A.COMB.pep: \*  
2: /cgn2\_6/ptodata/1/1aa/5B.COMB.pep: \*  
3: /cgn2\_6/ptodata/1/1aa/6A.COMB.pep: \*  
4: /cgn2\_6/ptodata/1/1aa/6B.COMB.pep: \*  
5: /cgn2\_6/ptodata/1/1aa/PCFUS.COMB.pep: \*  
6: /cgn2\_6/ptodata/1/1aa/backfilest.pep: \*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	117	8.6	342	US-09-270-767-32874	Sequence 32874, A
2	116	8.6	790	US-08-991-761A-13	Sequence 13, Appl
3	113	8.3	256	US-09-270-767-33129	Sequence 33129, A
4	113	8.3	256	US-09-270-767-48346	Sequence 48346, A
5	105	7.8	812	US-08-991-761A-7	Sequence 7, Appl
6	104	7.7	810	US-08-991-761A-11	Sequence 11, Appl
7	101.5	7.5	790	US-08-469-486-54	Sequence 54, Appl
8	101.5	7.5	790	US-08-469-658-54	Sequence 54, Appl
9	101	7.5	791	US-08-643-219-1	Sequence 1, Appl
10	101	7.5	791	US-09-131-995-1	Sequence 1, Appl
11	101	7.5	791	US-08-832-0878-1	Sequence 1, Appl
12	101	7.5	791	US-08-851-350-1	Sequence 1, Appl
13	101	7.5	791	US-09-132-154-1	Sequence 1, Appl
14	101	7.5	791	US-08-991-761A-6	Sequence 6, Appl
15	101	7.5	791	US-08-924-287A-1	Sequence 1, Appl
16	101	7.5	810	US-07-854-603-2	Sequence 2, Appl
17	101	7.5	810	US-08-147-000B-29	Sequence 29, Appl
18	101	7.5	810	US-09-086-514-1	Sequence 1, Appl
19	101	7.5	810	US-09-192-012-5	Sequence 5, Appl
20	101	7.5	810	US-09-403-736-1	Sequence 1, Appl
21	101	7.5	810	US-09-701-265-1	Sequence 1, Appl
22	101	7.5	810	5200340-8	Patent No. 5200340
23	101	7.5	810	5200340-8	Patent No. 5200340
24	101	7.5	814	US-08-750-711-1	Sequence 1, Appl
25	99	7.3	812	US-08-248-629A-1	Sequence 1, Appl
26	99	7.3	812	US-08-451-932-1	Sequence 1, Appl
27	99	7.3	812	US-08-452-260-1	Sequence 1, Appl

28	99	7.3	812	1	US-08-326-785-1	Sequence 1, Appl
29	99	7.3	812	2	US-08-612-788-1	Sequence 1, Appl
30	99	7.3	812	2	US-08-605-598B-1	Sequence 1, Appl
31	99	7.3	812	2	US-08-429-743-1	Sequence 1, Appl
32	99	7.3	812	2	US-08-866-735-1	Sequence 1, Appl
33	99	7.3	812	3	US-09-066-028-1	Sequence 1, Appl
34	99	7.3	812	3	US-09-192-012-3	Sequence 3, Appl
35	99	7.3	812	4	US-09-335-325-1	Sequence 1, Appl
36	99	7.3	812	4	US-08-991-761A-12	Sequence 12, Appl
37	99	7.3	812	5	PCT-US95-05107-1	Sequence 1, Appl
38	98.5	7.3	451	3	US-09-377-250-1	Sequence 1, Appl
39	98	7.2	452	3	US-09-377-250-4	Sequence 4, Appl
40	95	7.0	685	4	US-09-270-767-43374	Sequence 43374, A
41	93.5	6.9	915	4	US-09-907-794A-34	Sequence 34, Appl
42	93.5	6.9	915	4	US-09-905-125A-34	Sequence 34, Appl
43	93.5	6.9	915	4	US-09-902-775A-34	Sequence 34, Appl
44	93.5	6.9	915	4	US-09-906-700-34	Sequence 34, Appl
45	93.5	6.9	915	4	US-09-903-603A-34	Sequence 34, Appl

ALIGNMENTS

```
RESULT 1
US-09-270-767-32874
Sequence 32874, Application US/09270767
Patent No. 6703491
GENERAL INFORMATION:
APPLICANT: Homburger et al.
TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
FILE REFERENCE: File Reference: 7326-094
CURRENT APPLICATION NUMBER: US/09/270,767
CURRENT FILING DATE: 1999-03-17
NUMBER OF SEQ ID NOS: 62517
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 32874
LENGTH: 342
TYPE: PRT
ORGANISM: Drosophila melanogaster
US-09-270-767-32874

Query Match      8.6%, Score 117; DB 4; Length 342;
Best Local Similarity 23.6%, Pred. No. 0.00023;
Matches 64; Conservative 33; Mismatches 98; Indels 76; Gaps 15;

QY      3 KILVGFARFVVSADSVHR-----VHRCISACLNAFDTGFCBESVYYPVDAECTL- 50
DB      36 EILNGY---VLSAPSKMLDTPGLTLMULTCLEACON-----NESCNAVNY--ETGLCVLF 85
QY      51 -NTEDRLDRPDLFVDEHEDTYIYLDNNCAG---CECHHPDNFKTSGLINDQGFALIAQC 106
DB      86 KTYVDKLP--GSLSSQFPVFTTYAQKSCLGVRPSKAWCIDR----- 126
QY      107 YAPVTVQVAVEGRQLSDLDHSFEGLEISECELIQRLSVTANDFNCKSFMYNLTFRS 166
DB      127 -----VQGRLEPEHVSSQTVLSRDCLELC-----LGTEFPCRRANYRHSGL 171
QY      167 CVLSD-EBSRPLGRANLAIEVGMWTFESRGV--DS---FTRVPOMLIVGFASFVMEVNP 219
DB      172 CELSDMDRRLITLSAGSVSEVPYDGADYLENNCAEBSKLCFRIKSLIKTVDS--VHQDIN 230
QY      220 SYVMCLDQCTSPPEPTGQNFVCKSVMYYYNE 250
DB      231 TIDECRLCLNSP-----YKCHS--YDYND 253

RESULT 2
US-08-991-761A-13
Sequence 13, Application US/08991761A
Patent No. 6576609
GENERAL INFORMATION:
APPLICANT: Soff, Gerald
APPLICANT: Gately, Stephen
```

```

; APPLICANT: Twardowski, Przemyslaw
; TITLE OF INVENTION: "Methods and Compositions for Generating
; TITLE OF INVENTION: Angiostatin"
; NUMBER OF SEQUENCES: 16
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Sheridan Ross P.C.
; STREET: 1700 Lincoln St., Suite 3500
; CITY: Denver
; STATE: CO
; COUNTRY: USA
; ZIP: 80203
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/991,761A
; FILING DATE:
; CLASSIFICATION: 1642
; ATTORNEY/AGENT INFORMATION:
; NAME: Crook, Wannell M.
; REGISTRATION NUMBER: 31,071
; REFERENCE/DOCKET NUMBER: 3501-16-1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (303) 863-8700
; TELEFAX: (303) 863-0223
; INFORMATION FOR SEQ ID NO: 13:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 790 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-991-761A-13

```

```

Query Match      8.6%; Score 116; DB 4; Length 790;
Best Local Similarity 19.5%; Pred. No. 0.00098;
Matches 55; Conservative 43; Mismatches 100; Indels 84; Gaps 12;

```

```

QY 5 LVGFARFVVSADSVHRCLSACINAFDTFGFEGESVMYYPVDAECIINTEDRLDRPLFVD 64
| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 14 LFLSLRQVAVARSVESCAKAC--EAST-NFICRAFOYHSHDOOCVMAENSKTSP--1A 67
| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 65 EHEDTVYIYLDNNCAGCECHWHFDNFKTSGILNDQGFALAAQCYAPVYQYVAVEG 119
| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 68 RMDVVLFEKRIYLSSECKTGNGKNGYGTSTKTSVYICQMSVSSPHIPIKTSPEKPLAG 127
| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 120 RQ-----LSDELHSPFEGLELSECELCCTO-----RLSVTANDFNC 155
| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 128 LEENYCRNPNDKRGPMCYTTPDETRPDYCDIPECEDECHGSGEHYEGKISTKMSGIGC 187
| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 156 KS-----FKYSNLTFRS-CVLSDRSRPLGRANLAIEVPGWTFESRGVPSF 199
| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 188 QSWGSGSPHAGYLPKSFPMKNLKNMYCRNPDEPRP-----W-----CF 227
| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 200 TRVPQMLLVGFASFVMEVNPVSMCLDQCTSPPEPTGQFNC 241
| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 228 TTDPM-----KRWECDFP-----KCTTPPTSGFTYQC 256
| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

```

```

RESULT 3
US-09-270-767-33129
; Sequence 33129 Application US/09270767
; Patent No. 6703491
; GENERAL INFORMATION:
; APPLICANT: Homburger et al.
; TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
; FILE REFERENCE: File Reference: 7326-094
; CURRENT APPLICATION NUMBER: US/09/270,767
; CURRENT FILING DATE: 1999-03-17
; NUMBER OF SEQ ID NOS: 62517
; SOFTWARE: Patentin Ver. 2.0

```

```

; SEQ ID NO 33129
; LENGTH: 256
; TYPE: PRT
; ORGANISM: Drosophila melanogaster
; US-09-270-767-33129

```

```

Query Match      8.3%; Score 113; DB 4; Length 256;
Best Local Similarity 21.3%; Pred. No. 0.00041;
Matches 47; Conservative 28; Mismatches 82; Indels 64; Gaps 7;

```

```

QY 3 KLVGFARFVVSADSVHRCLSACINAFDTFGFEGESVMYYPVDAECIINTEDRLDRPLF 62
| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 47 KEIRGLDNALITYSTKEACLSACLN---ERRFVCRSVYDYNNMKCVLSDDRSGGFV 103
| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 63 VDEHEDTVIYLDNNC---AGCECHWHFDNFKTSGILNDQGFALAAQCYAPVYQYVAVE 118
| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 104 QLVDAQGTDFENLCLKPAQCKNNRSFGNSQKMGVSEK-----VAQYVGL- 150
| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 119 GRQLSDELHSPFEGLELSECELCCTQRLSVTANDFNCKSFMYSNLTR----- 165
| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 151 -----HYTDELQVTSBSACRLACEISEFLCGRFYLGGPQGSQYNCRLYHHDH 201
| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 166 -----SCVLSDER-----SRPLGRANLAIEVPGWTFESR 194
| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 202 KTLPDGPSTYLNHERPLIDHGEPITQ-----YFENQ 232
| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

```

```

RESULT 4
US-09-270-767-48346
; Sequence 48346, Application US/09270767
; Patent No. 6703491
; GENERAL INFORMATION:
; APPLICANT: Homburger et al.
; TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
; FILE REFERENCE: File Reference: 7326-094
; CURRENT APPLICATION NUMBER: US/09/270,767
; NUMBER OF SEQ ID NOS: 62517
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 48346
; LENGTH: 256
; TYPE: PRT
; ORGANISM: Drosophila melanogaster
; US-09-270-767-48346

```

```

Query Match      8.3%; Score 113; DB 4; Length 256;
Best Local Similarity 21.3%; Pred. No. 0.00041;
Matches 47; Conservative 28; Mismatches 82; Indels 64; Gaps 7;

```

```

QY 3 KLVGFARFVVSADSVHRCLSACINAFDTFGFEGESVMYYPVDAECIINTEDRLDRPLF 62
| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 47 KEIRGLDNALITYSTKEACLSACLN---ERRFVCRSVYDYNNMKCVLSDDRSGGFV 103
| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 63 VDEHEDTVIYLDNNC---AGCECHWHFDNFKTSGILNDQGFALAAQCYAPVYQYVAVE 118
| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 104 QLVDAQGTDFENLCLKPAQCKNNRSFGNSQKMGVSEK-----VAQYVGL- 150
| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 119 GRQLSDELHSPFEGLELSECELCCTQRLSVTANDFNCKSFMYSNLTR----- 165
| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 151 -----HYTDELQVTSBSACRLACEISEFLCGRFYLGGPQGSQYNCRLYHHDH 201
| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 166 -----SCVLSDER-----SRPLGRANLAIEVPGWTFESR 194
| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 202 KTLPDGPSTYLNHERPLIDHGEPITQ-----YFENQ 232
| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

```

```

RESULT 5
US-08-991-761A-7
; Sequence 7, Application US/08991761A
; Patent No. 6576609
; GENERAL INFORMATION:
; APPLICANT: Soft, Gerald
; APPLICANT: Gately, Stephen

```

```

; APPLICANT: Twardowski, Przemyslaw
; TITLE OF INVENTION: "Methods and Compositions for Generating
; TITLE OF INVENTION: Angiostatin"
; NUMBER OF SEQUENCES: 16
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Sheridan Ross P.C.
; STREET: 1700 Lincoln St., Suite 3500
; CITY: Denver
; STATE: CO
; COUNTRY: USA
; ZIP: 80203
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/991,761A
; FILING DATE:
; CLASSIFICATION: 1642
; ATTORNEY/AGENT INFORMATION:
; NAME: Crook, Marnell M.
; REGISTRATION NUMBER: 31,071
; REFERENCE/DOCKET NUMBER: 3501-16-1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (303) 863-9700
; TELEFAX: (303) 863-0223
; INFORMATION FOR SEQ ID NO: 7:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 812 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-991-761A-7

```

```

Query Match 7.8%; Score 105; DB 4; Length 812;
Best Local Similarity 18.3%; Pred. No. 0.017;
Matches 55; Conservative 42; Mismatches 83; Indels 120; Gaps 17;

```

```

QY 5 LVGFAREVVSADSVHRCISACINAFDTGFECESVMYYPVDAECILNTEDLRDLFVD 64
DB 40 LLSIRKMLAGRSVDECAKCEETD--FVCRARQYHSKEQCQVMAENSGNTP---VF 93
QY 65 EHEDTVIYLDN---NC-----ACCECH-W-----HEDNFKTSLINDQ 98
DB 94 RMRDVLVEKRIYLLCKTGNGQYRGTTAETKSGVTCQKWSATSPHVPKF-----SPE 147
QY 99 QFALNA---OCYAP-----YVQYVAVEGRQLSDELDPFEGLELSECEELCTQ- 144
DB 148 KFPPLGLEENYCRNPDNDENGPMCTT-----DPDKRYDCDIFECEDCKMHC 195
QY 145 -----RLSVYANDFNCKS-----FMYSNLTFRS-CVLSDESRRLPGRAN 181
DB 196 SGENYEGKIATKMSGRDCQAMDSQSPHAGYIPSKFPMNKLKKNYCRNPDGSRP- 250
QY 182 LAEVGWTYFESRGVPSFTRPQMLLVGFASFMENVSVMCLDOCTSPPETQNFVC 241
DB 251 -----W-----CFTTDPQ-----KRWEFCDIP-----RCTTPPSSGPTYOC 282

```

```

RESULT 6
US-08-991-761A-11
; Sequence 11, Application US/08991761A
; Patent No. 6576609
; GENERAL INFORMATION:
; APPLICANT: Self, Gerald
; APPLICANT: Gately, Stephen
; APPLICANT: Twardowski, Przemyslaw
; TITLE OF INVENTION: "Methods and Compositions for Generating
; TITLE OF INVENTION: Angiostatin"
; NUMBER OF SEQUENCES: 16
; CORRESPONDENCE ADDRESS:

```

```

; ADDRESSEE: Sheridan Ross P.C.
; STREET: 1700 Lincoln St., Suite 3500
; CITY: Denver
; STATE: CO
; COUNTRY: USA
; ZIP: 80203
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/991,761A
; FILING DATE:
; CLASSIFICATION: 1642
; ATTORNEY/AGENT INFORMATION:
; NAME: Crook, Marnell M.
; REGISTRATION NUMBER: 31,071
; REFERENCE/DOCKET NUMBER: 3501-16-1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (303) 863-9700
; TELEFAX: (303) 863-0223
; INFORMATION FOR SEQ ID NO: 11:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 810 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-991-761A-11

```

```

Query Match 7.7%; Score 104; DB 4; Length 810;
Best Local Similarity 18.1%; Pred. No. 0.021;
Matches 54; Conservative 38; Mismatches 89; Indels 118; Gaps 14;

```

```

QY 5 LVGFAREVVSADSVHRCISACINAFDTGFECESVMYYPVDAECILNTEDLRDLFVD 64
DB 33 LFSITKKQLGAGSIECAKCEEBE--FVCRARQYHSKEQCQVMAENSGNTP---VF 86
QY 65 EHEDTVIYLDNKNAGCEBHMFPDNFKTS-----GIINDQFALNAOCY---APVITQY 115
DB 87 RMRDVLVEF-----EKVYLSBCKTGNGKNGYRGYTSKRTGTGTCQKWSATSPH 133
QY 116 AVEGRQLSDELDPFEGLE-----LSRCEELCTQ- 144
DB 134 ---RPTSPATHSBGLSENYCRNPDNDGSGPMCTTDPBERPDYCDIFECEDCKMHC 189
QY 145 -----RLSVYANDFNCKS-----FMYSNLTFRS-CVLSDESRRLPGRANL 182
DB 190 GENYDGKISKTMSGLBQAMDSQSPHAGYIPSKFPMNKLKKNYCRNPDGSRP- 243
QY 183 AEPGWTYFESRGVPSFTRPQMLLVGFASFMENVSVMCLDOCTSPPETQNFVC 241
DB 244 -----W-----CFTTDPN-----KRWELCDIP-----RCTTPPSSGPTYOC 275

```

```

RESULT 7
US-08-469-486-54
; Sequence 54, Application US/08469486
; Patent No. 5739281
; GENERAL INFORMATION:
; APPLICANT: Thogersen, Hans Christian
; APPLICANT: Holsted, Thor Ias
; APPLICANT: Etzerodt, Michael
; TITLE OF INVENTION: Improved method for the refolding of
; TITLE OF INVENTION: proteins
; NUMBER OF SEQUENCES: 58
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson
; STREET: 225 Franklin Street
; CITY: Boston
; STATE: Massachusetts
; COUNTRY: USA

```





SOFTWARE: FastSeq Version 2.0  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/832,087B  
FILING DATE: 03-APR-1997  
CLASSIFICATION: 514  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/643,219  
FILING DATE: 06-MAY-1996  
ATTORNEY/AGENT INFORMATION:  
NAME: Casuto, Dianne  
REGISTRATION NUMBER: 40,943  
REFERENCE/DOCKET NUMBER: 5940.US.P1  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 847-938-3137  
TELEFAX: 847-938-2623  
TELEX:  
INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 791 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE:  
DESCRIPTION: protein  
US-08-832-087B-1

Query Match 7.5%; Score 101; DB 2; Length 791;  
Best Local Similarity 18.0%; Pred. No. 0.045;  
Matches 53; Conservative 42; Mismatches 90; Indels 110; Gaps 13;

QY 5 LVGFAREVVSADSVHRCISACLNAPDTFGECESVMYYPVDAECILNTEDR-----LDRP 59  
DB 14 LFSVTKKQJGAGSIECAAKCEDEE---FTCRAPQYHSKEQCVCYMAENRSSIILIMR 70  
QY 60 DLFDVDEHEDVIVYLDNNCAGCECHWHFDFNFKTSGILNDQOFAIAQCYAPVYTYVAVEG 119  
DB 71 DVVLFPEKK---VYLSECKTGNGKNYGTMSKTXNGITCQKMSST---SPH----- 114  
QY 120 RQLSDELHSPFGLF-----LSECEELCTQ----- 144  
DB 115 RPRFSPATHPSGGLSENYCNPNDPQGPWCYTTDPEKRYDCDILCEEEBCMHGSGENY 174  
QY 145 --RLSTVANDFNCKS-----FMYSNLTRS-CVLSDESRPLGRANILAEVP 186  
DB 175 DGKISTKMSGLECQAMDQSPRAHGYIPSKFPKNKLNKXNCRNPDBELRP----- 224  
QY 187 GWTYFESRGVPSFTVPQMLVGFASFVMEVNPVSTMCLDOCTSPPETGQNFVC 241  
DB 225 -W-----CFTTDPN-----KRWELCDIP-----RCTTPPSGSPTYQC 256

RESULT 12  
US-08-851-350-1  
Sequence 1, Application US/08851350  
Patent No. 6057122  
GENERAL INFORMATION:  
APPLICANT: Abbott Laboratories  
TITLE OF INVENTION: NOVEL ANTIANGIOGENIC PEPTIDES,  
TITLE OF INVENTION: POLYNUCLEOTIDES ENCODING SAME AND METHODS  
NUMBER OF SEQUENCES: 38  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Abbott Laboratories  
STREET: 100 Abbott Park Road  
CITY: Abbott Park  
STATE: IL  
COUNTRY: USA  
ZIP: 60064-3500  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: DOS  
SOFTWARE: FastSeq Version 2.0  
CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/851,350  
FILING DATE: 05-MAY-1997  
CLASSIFICATION: 514  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER:  
FILING DATE:  
ATTORNEY/AGENT INFORMATION:  
NAME: Casuto, Dianne  
REGISTRATION NUMBER: 40,943  
REFERENCE/DOCKET NUMBER: 5940.US.P2  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 847-938-3137  
TELEFAX: 847-938-2623  
TELEX:  
INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 791 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE:  
DESCRIPTION: protein  
US-08-851-350-1

Query Match 7.5%; Score 101; DB 3; Length 791;  
Best Local Similarity 18.0%; Pred. No. 0.045;  
Matches 53; Conservative 42; Mismatches 90; Indels 110; Gaps 13;

QY 5 LVGFAREVVSADSVHRCISACLNAPDTFGECESVMYYPVDAECILNTEDR-----LDRP 59  
DB 14 LFSVTKKQJGAGSIECAAKCEDEE---FTCRAPQYHSKEQCVCYMAENRSSIILIMR 70  
QY 60 DLFDVDEHEDVIVYLDNNCAGCECHWHFDFNFKTSGILNDQOFAIAQCYAPVYTYVAVEG 119  
DB 71 DVVLFPEKK---VYLSECKTGNGKNYGTMSKTXNGITCQKMSST---SPH----- 114  
QY 120 RQLSDELHSPFGLF-----LSECEELCTQ----- 144  
DB 115 RPRFSPATHPSGGLSENYCNPNDPQGPWCYTTDPEKRYDCDILCEEEBCMHGSGENY 174  
QY 145 --RLSTVANDFNCKS-----FMYSNLTRS-CVLSDESRPLGRANILAEVP 186  
DB 175 DGKISTKMSGLECQAMDQSPRAHGYIPSKFPKNKLNKXNCRNPDBELRP----- 224  
QY 187 GWTYFESRGVPSFTVPQMLVGFASFVMEVNPVSTMCLDOCTSPPETGQNFVC 241  
DB 225 -W-----CFTTDPN-----KRWELCDIP-----RCTTPPSGSPTYQC 256

RESULT 13  
US-09-132-154-1  
Sequence 1, Application US/09132154  
Patent No. 6251867  
GENERAL INFORMATION:  
APPLICANT: Davidson, Donald J.  
TITLE OF INVENTION: NOVEL ANTIANGIOGENIC PEPTIDES  
TITLE OF INVENTION: AND METHODS FOR INHIBITING ANGIOGENESIS  
NUMBER OF SEQUENCES: 14  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Abbott Laboratories  
STREET: 100 Abbott Park Road  
CITY: Abbott Park  
STATE: IL  
COUNTRY: USA  
ZIP: 60064  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: DOS  
SOFTWARE: FastSeq Version 2.0  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/132,154  
FILING DATE:  
CLASSIFICATION:



PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US/08/832,087  
FILING DATE: 03-APR-1997  
APPLICATION NUMBER: 08/543,219  
FILING DATE: 06-MAY-1996  
ATTORNEY/AGENT INFORMATION:  
NAME: Casuto, Dianne  
REGISTRATION NUMBER: 40,943  
REFERENCE/DOCKET NUMBER: 5940.US.P1  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 847-938-3137  
TELEFAX: 847-938-2623  
TELEX:  
INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 791 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE:  
DESCRIPTION: protein  
US-09-132-154-1

Query Match 7.5%; Score 101; DB 3; Length 791;  
Best Local Similarity 18.0%; Pred. No. 0.045;  
Matches 53; Conservative 42; Mismatches 90; Indels 110; Gaps 13;

QY 5 LVGFARVNSADSVHRCISACLNAPDTGFGECSVMYYPVDAECILNTEDR-----LDRP 59  
DB 14 LFSVTKQLGAGSIEBCAKCEDEE---FTCRAPQYHSKEQCVIMANKSIIIMR 70  
QY 60 DLVDEHEDTVIYLDNNCAGECECHWFNFKTSGLINDQOFAIAQCYAPYVTOYVAVEG 119  
DB 71 DVLVFEKK---VYLSSECTGNGKNYRGMTSKTKNGITQKKMST---SPH----- 114  
QY 120 RQLSDELHDSFEGL------LSCEBELCTQ----- 144  
DB 115 RPRFSPATHPSEGLEENYCRPNDDPGPWCYTTDPEKRYDYCDILECEEEBCHGSGENY 174  
QY 145 --RUSVTANDFNCKS-----FMYSNLTRS--CVLSDERBRPLGRANLAVRP 186  
DB 175 DGLSKTWSGLECCAMWSQSPHAGYIPSKFPNKNLKNYCRNPREDLRP----- 224

QY 187 GMTYFESRGVPSFTVRPQMLVGFASFVMEVNPVSTMCLDOCTSPPEPTGQNFVC 241  
DB 225 -W-----CFTTDPN-----KRWELCDIP-----RCTTPPSSGPTGYQC 256

RESULT 14  
US-08-991-761A-6  
Sequence 6, Application US/08991761A  
Patent No. 6576609  
GENERAL INFORMATION:  
APPLICANT: Sofi, Gerald  
APPLICANT: Gately, Stephen  
APPLICANT: Twardowski, Przemyslaw  
TITLE OF INVENTION: "Methods and Compositions for Generating  
TITLE OF INVENTION: Angiostatin"  
NUMBER OF SEQUENCES: 16  
CORRESPONDENCE ADDRESSES:  
ADDRESSER: Sheridan Rose P.C.  
STREET: 1700 Lincoln St., Suite 3500  
CITY: Denver  
STATE: CO  
COUNTRY: USA  
ZIP: 80203  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/991,761A  
FILING DATE:

CLASSIFICATION: 1642  
ATTORNEY/AGENT INFORMATION:  
NAME: Crook, Marnell M.  
REGISTRATION NUMBER: 31,071  
REFERENCE/DOCKET NUMBER: 3501-16-1  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (303) 863-9700  
TELEFAX: (303) 863-0223  
INFORMATION FOR SEQ ID NO: 6:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 791 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-991-761A-6

Query Match 7.5%; Score 101; DB 4; Length 791;  
Best Local Similarity 18.0%; Pred. No. 0.045;  
Matches 53; Conservative 42; Mismatches 90; Indels 110; Gaps 13;

QY 5 LVGFARVNSADSVHRCISACLNAPDTGFGECSVMYYPVDAECILNTEDR-----LDRP 59  
DB 14 LFSVTKQLGAGSIEBCAKCEDEE---FTCRAPQYHSKEQCVIMANKSIIIMR 70  
QY 60 DLVDEHEDTVIYLDNNCAGECECHWFNFKTSGLINDQOFAIAQCYAPYVTOYVAVEG 119  
DB 71 DVLVFEKK---VYLSSECTGNGKNYRGMTSKTKNGITQKKMST---SPH----- 114  
QY 120 RQLSDELHDSFEGL------LSCEBELCTQ----- 144  
DB 115 RPRFSPATHPSEGLEENYCRPNDDPGPWCYTTDPEKRYDYCDILECEEEBCHGSGENY 174  
QY 145 --RUSVTANDFNCKS-----FMYSNLTRS--CVLSDERBRPLGRANLAVRP 186  
DB 175 DGLSKTWSGLECCAMWSQSPHAGYIPSKFPNKNLKNYCRNPREDLRP----- 224

QY 187 GMTYFESRGVPSFTVRPQMLVGFASFVMEVNPVSTMCLDOCTSPPEPTGQNFVC 241  
DB 225 -W-----CFTTDPN-----KRWELCDIP-----RCTTPPSSGPTGYQC 256

RESULT 15  
US-08-924-287A-1  
Sequence 1, Application US/08924287A  
Patent No. 6699838  
GENERAL INFORMATION:  
APPLICANT: Abbott Laboratories  
APPLICANT: Davidson, Donald J.  
TITLE OF INVENTION: NOVEL ANTIANGIOGENIC PEPTIDES,  
TITLE OF INVENTION: POLYNUCLEOTIDES ENCODING SAME AND METHODS FOR INHIBITING  
TITLE OF INVENTION: ANGIOGENESIS  
FILE REFERENCE: 5940.US.P3  
CURRENT APPLICATION NUMBER: US/08/924,287A  
CURRENT FILING DATE: 1997-09-05  
PRIOR APPLICATION NUMBER: US 08/851,350  
PRIOR FILING DATE: 1997-05-05  
PRIOR APPLICATION NUMBER: US 08/832,087  
PRIOR FILING DATE: 1997-04-03  
PRIOR APPLICATION NUMBER: US 08/643,219  
PRIOR FILING DATE: 1996-05-03  
NUMBER OF SEQ ID NOS: 40  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 1  
LENGTH: 791  
TYPE: PRT  
ORGANISM: Homo sapiens  
US-08-924-287A-1

Query Match 7.5%; Score 101; DB 4; Length 791;  
Best Local Similarity 18.0%; Pred. No. 0.045;  
Matches 53; Conservative 42; Mismatches 90; Indels 110; Gaps 13;

```

QY 5 LVGPAREVVSADSVHRCLSACLNAPDTGEGECESVMYYPVDAECILNTEDR----LDRP 59
Db 14 LFSVTKKQJLGAGSIEECACACEDEE---FTCRAFOYHSKEOCVIMAEHRKSSIIIRWR 70
QY 60 DLFPVDEHEDVIYLDNNCAGCECHWHEDNFKTSGIINDQOFALAAOCYAPYTOYAVBEG 119
Db 71 DVVLPFEKK--VYLSBCKTGNKNGKGTMSKTYNGITCQKMSST---SPH----- 114
QY 120 RQLSDELDHSFEGILE-----LSCEBELCTQ----- 144
Db 115 RPRFSPATHSEGLSENYCNPNDNDPOGFWCYTTDPEKRYDYCDILECEHECHGSGENY 174
QY 145 --RLSVTANDFNCKS-----FMTSNLTRS-CVLSDESRPLGRANLAEPV 186
Db 175 DGKISKTMSGLECOAMDQSPHAGYIPSKFPNKNLKKNYCRNPDRELRP----- 224
QY 187 GWTFPESRGVPSFTRVPOMLLVGFASFVMENVPSTMCLDQCTSPPETGQNFVC 241
Db 225 -W-----CFTTDPN-----KRWELCDIP-----RCTTPPSSGPTYQC 256

```

Search completed: March 31, 2005, 02:07:44  
Job time : 44 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OW protein - protein search, using sw model

Run on: March 31, 2005, 02:00:10 ; Search time 140 Seconds  
(without alignments)  
591.252 Million cell updates/sec

Title: US-10-771-708-11

Perfect score: 1354  
Sequence: 1 EQLIVGFARFVNSADSVHR.....PPETGQNFVCKSVWYVYNE 250

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 1407402 seqs, 331100923 residues

Total number of hits satisfying chosen parameters: 1407402

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :

Published Applications AA:\*

- 1: /cgn2\_6/ptodata/2/pubpaa/US07\_PUBCOMB.pep.\*
- 2: /cgn2\_6/ptodata/2/pubpaa/PCT\_NEW\_PUB.pep.\*
- 3: /cgn2\_6/ptodata/2/pubpaa/US06\_NEW\_PUB.pep.\*
- 4: /cgn2\_6/ptodata/2/pubpaa/US07\_NEW\_PUB.pep.\*
- 5: /cgn2\_6/ptodata/2/pubpaa/PCTUS\_PUBCOMB.pep.\*
- 6: /cgn2\_6/ptodata/2/pubpaa/US08\_NEW\_PUB.pep.\*
- 7: /cgn2\_6/ptodata/2/pubpaa/US08\_PUBCOMB.pep.\*
- 8: /cgn2\_6/ptodata/2/pubpaa/US08\_PUBCOMB.pep.\*
- 9: /cgn2\_6/ptodata/2/pubpaa/US09A\_PUBCOMB.pep.\*
- 10: /cgn2\_6/ptodata/2/pubpaa/US09B\_PUBCOMB.pep.\*
- 11: /cgn2\_6/ptodata/2/pubpaa/US09C\_PUBCOMB.pep.\*
- 12: /cgn2\_6/ptodata/2/pubpaa/US09\_NEW\_PUB.pep.\*
- 13: /cgn2\_6/ptodata/2/pubpaa/US10A\_PUBCOMB.pep.\*
- 14: /cgn2\_6/ptodata/2/pubpaa/US10B\_PUBCOMB.pep.\*
- 15: /cgn2\_6/ptodata/2/pubpaa/US10C\_PUBCOMB.pep.\*
- 16: /cgn2\_6/ptodata/2/pubpaa/US10D\_PUBCOMB.pep.\*
- 17: /cgn2\_6/ptodata/2/pubpaa/US10E\_PUBCOMB.pep.\*
- 18: /cgn2\_6/ptodata/2/pubpaa/US11\_NEW\_PUB.pep.\*
- 19: /cgn2\_6/ptodata/2/pubpaa/US60\_NEW\_PUB.pep.\*
- 20: /cgn2\_6/ptodata/2/pubpaa/US60\_PUBCOMB.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	116	8.6	790	US-09-825-751A-70	Sequence 70, Appl
2	105	7.8	812	US-09-825-751A-71	Sequence 71, Appl
3	101	7.5	567	US-10-741-601-413	Sequence 413, App
4	101	7.5	569	US-09-946-893-5	Sequence 5, Appl1
5	101	7.5	576	US-09-946-893-8	Sequence 8, Appl1
6	101	7.5	576	US-09-946-893-6	Sequence 6, Appl1
7	101	7.5	791	US-09-967-386-1	Sequence 1, Appl1
8	101	7.5	791	US-10-304-287-1	Sequence 1, Appl1
9	101	7.5	791	US-10-360-101-257	Sequence 257, App
10	101	7.5	791	US-10-778-423-1	Sequence 1, Appl1
11	101	7.5	791	US-10-753-646-1	Sequence 1, Appl1
12	101	7.5	791	US-10-735-577-1	Sequence 1, Appl1
13	101	7.5	810	US-09-946-893-2	Sequence 2, Appl1

14	101	7.5	810	14	US-10-193-656-2	Sequence 2, Appl1
15	101	7.5	810	14	US-10-237-144-1	Sequence 1, Appl1
16	101	7.5	810	15	US-10-135-8728-4	Sequence 4, Appl1
17	101	7.5	810	15	US-10-450-976-2	Sequence 2, Appl1
18	101	7.5	810	15	US-10-415-012-4	Sequence 4, Appl1
19	101	7.5	810	16	US-10-741-601-409	Sequence 140, App
20	101	7.5	810	16	US-10-872-198-140	Sequence 6, Appl1
21	99	7.3	333	15	US-10-135-8728-6	Sequence 2, Appl1
22	99	7.3	466	16	US-10-753-164-2	Sequence 1, Appl1
23	99	7.3	812	9	US-09-788-142-1	Sequence 1, Appl1
24	99	7.3	812	9	US-09-761-120-1	Sequence 1, Appl1
25	99	7.3	812	9	US-09-873-676-81	Sequence 81, Appl1
26	99	7.3	812	9	US-09-335-325-1	Sequence 1, Appl1
27	99	7.3	812	14	US-10-131-241-1	Sequence 1, Appl1
28	99	7.3	812	14	US-10-127-066-1	Sequence 1, Appl1
29	99	7.3	812	15	US-10-402-364-1	Sequence 1, Appl1
30	99	7.3	812	15	US-10-401-108-1	Sequence 1, Appl1
31	98	7.2	608	13	US-10-087-192-795	Sequence 795, App
32	95	7.0	548	15	US-10-369-493-22599	Sequence 22599, A
33	93.5	6.9	741	16	US-10-408-765A-1187	Sequence 1187, Ap
34	93.5	6.9	794	11	US-09-833-245-1292	Sequence 1292, Ap
35	93.5	6.9	811	15	US-10-363-616-394	Sequence 394, App
36	93.5	6.9	830	15	US-10-353-616-393	Sequence 393, App
37	93.5	6.9	896	15	US-10-004-378A-16	Sequence 16, Appl
38	93.5	6.9	896	15	US-10-428-275-158	Sequence 158, App
39	93.5	6.9	914	15	US-10-428-275-156	Sequence 156, App
40	93.5	6.9	915	9	US-09-909-330-34	Sequence 34, Appl
41	93.5	6.9	915	9	US-09-909-088B-34	Sequence 34, Appl
42	93.5	6.9	915	9	US-09-905-291A-34	Sequence 34, Appl
43	93.5	6.9	915	9	US-09-902-853-34	Sequence 34, Appl
44	93.5	6.9	915	9	US-09-907-824-34	Sequence 34, Appl
45	93.5	6.9	915	9	US-09-907-841-34	Sequence 34, Appl

ALIGNMENTS

RESULT 1  
US-09-825-751A-70  
; Sequence 70, Application US/09825751A  
; Publication No. US20030065140A1  
; GENERAL INFORMATION:  
; APPLICANT: Curagen Corporation  
; APPLICANT: Vernet, Corine A.M.  
; APPLICANT: Fernandes, Elma R.  
; APPLICANT: Taupier, Raymond J  
; APPLICANT: Quinn, Kerry E  
; APPLICANT: Spytek, Kimberly A  
; APPLICANT: Raetelli, Luca  
; APPLICANT: Herrman, John L.  
; TITLE OF INVENTION: Novel Proteins and Nucleic Acids Encoding Same  
; FILE REFERENCE: 15966-750  
; CURRENT APPLICATION NUMBER: US/09/825,751A  
; PRIOR FILING DATE: 2001-04-30  
; PRIOR APPLICATION NUMBER: 60/194,314  
; PRIOR FILING DATE: 2000-04-03  
; PRIOR APPLICATION NUMBER: 60/225,693  
; PRIOR FILING DATE: 2000-08-16  
; NUMBER OF SEQ ID NOS: 85  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 70  
; LENGTH: 790  
; TYPE: PRT  
; ORGANISM: Sus scrofa  
; US-09-825-751A-70  
Query Match 8.6%; Score 116; DB 10; Length 790;  
Best Local Similarity 19.5%; Pred. No. 0.0064;  
Matches 55; Conservative 43; Mismatches 100; Indels 84; Gaps 12;  
QY 5 LVGFARFVNSADSVHRCTASCTNAFDTFGFECESVWYVVDACITLNTEDRLDRDPLVD 64  
DB 14 LFSLRKQVARSVECAKAC--EAET-NFICRAFGYHSKQOCVMAENSKTSP---IA 67

OY 65 EHEDTV-----IYLDNNACGCECHWFHDFNFKTSGIINDQOFAIAAOCTAPVYVAVG 119  
Db 68 RMDVVLFEKRIYLTSECKTGNKNGYGTTSKTSKVYICQKMSVSPHPIKYSSEKPLAG 127  
OY 120 RQ-----ISDELHSPFEGLESECELTQ-----RLSTANDFNC 155  
Db 128 LEBENYCRNPNDKRGKFWCYTDTDETRFDYCDIPECEBCHGSGEHYBKGISKTMSGIEC 187  
OY 156 KS-----FMYSNLTRS-CVLSDESRPLGRANLAEPVGTYPESRGVPSF 199  
Db 188 QSWGSGSPHAGVLPBKFPRKMLKMYCRNPDEPRP-----W-----CF 227  
OY 200 TRVPQMLLVGFASFMVENVPSTVMTCLDQCTSPPETGONFVC 241  
Db 228 TTDPN-----KRMFCDIP-----RCTTPPSTSGPYQC 256

RESULT 2  
US-09-825-751A-71  
; Sequence 71, Application US/09825751A  
; Publication No. US20030065140A1  
; GENERAL INFORMATION:  
; APPLICANT: Curagen Corporation  
; APPLICANT: Vermet, Corine A.M.  
; APPLICANT: Fernandes, Elma R.  
; APPLICANT: Taupier, Raymond J.  
; APPLICANT: Quinn, Kerry E.  
; APPLICANT: Spytek, Kimberly A.  
; APPLICANT: Rastelli, Luca  
; APPLICANT: Herrman, John L.  
; TITLE OF INVENTION: Novel Proteins and Nucleic Acids Encoding Same  
; FILE REFERENCE: 15966-750  
; CURRENT APPLICATION NUMBER: US/09/825,751A  
; CURRENT FILING DATE: 2001-04-30  
; PRIOR FILING DATE: 2000-04-03  
; PRIOR APPLICATION NUMBER: 60/194,314  
; PRIOR FILING DATE: 2000-04-03  
; PRIOR APPLICATION NUMBER: 60/225,693  
; PRIOR FILING DATE: 2000-08-16  
; NUMBER OF SEQ ID NOS: 85  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 71  
; LENGTH: 812  
; TYPE: PRT  
; ORGANISM: Bos taurus  
US-09-825-751A-71

Query Match 7.8%; Score 105; DB 10; Length 812;  
Best Local Similarity 18.3%; Pred. No. 0.096;  
Matches 55; Conservative 42; Mismatches 83; Indels 120; Gaps 17;  
OY 5 LVGFARVVSADSVHRCISACLNAPDTFGFECESVMYYPVDAECTINTEDRLDRPLDFVD 64  
Db 40 LLSLSKMLAGSVEDECAKCEBETD---FVCHAFQYHSKEQOCVMAENKSNTP---VF 93  
OY 65 EHEDTVYLDN---NC-----AGCECH-W-----HFDNFKTSGIINDQ 98  
Db 94 RMDVVLFEKRIYLTSECKTGNKNGYGTTAETKSGVTCQKMSATSPHVKF-----SPE 147  
OY 99 QFAIAA---QCYAP-----YTYVVAVGRQLSDELHSPFEGLESECELTQ- 144  
Db 148 KFFPLAGLEBNYCRNPNDENGKFWCYT-----DPDKRYVYCDIPREDECHMC 195  
OY 145 -----RLSTANDFNCKS-----FMYSNLTRS-CVLSDESRPLGRAN 181  
Db 196 SGENYSGKIATKMSGDCQAMDSQSPHAGYIPSKRPKNLKNYCRNPDEPRP----- 250  
OY 182 LAEVPQMTYFESRGVPSFTVPQMLLVGFASFMVENVPSTVMTCLDQCTSPPETGONFVC 241  
Db 251 -----W-----CFTTDPQ-----KRMFCDIP-----RCTTPPSSSGPYQC 282

RESULT 3

US-10-741-601-413  
; Sequence 413, Application US/10741601  
; Publication No. US20040166519A1  
; GENERAL INFORMATION:  
; APPLICANT: CARGILL, Michele et al.  
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH  
; FILE REFERENCE: CL001500  
; CURRENT APPLICATION NUMBER: US/10/741,601  
; CURRENT FILING DATE: 2003-12-22  
; NUMBER OF SEQ ID NOS: 26415  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 413  
; LENGTH: 567  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-10-741-601-413

Query Match 7.5%; Score 101; DB 16; Length 567;  
Best Local Similarity 18.0%; Pred. No. 0.16;  
Matches 53; Conservative 42; Mismatches 90; Indels 110; Gaps 13;  
OY 5 LVGFARVVSADSVHRCISACLNAPDTFGFECESVMYYPVDAECTINTEDR-----LDRP 59  
Db 33 LFSVTYKQLGAGSIECAKCEBDE---FTCRAFQYHSKEQOCVMAENKSSITIIIMR 89  
OY 60 DLFDVEHEDTVYIYLDNNACGCECHWFHDFNFKTSGIINDQOFAIAAOCTAPVYVAVG 119  
Db 90 DVVLFEK-----VYLSECKTGNKNGYGTTSKTSKVYICQKMSVSPHPIKYSSEKPLAG 133  
OY 120 RQSDDELHSPFEGLE-----LSECELTQ----- 144  
Db 134 RPRFSPATPSPRGLEBNYCRNPNDPQGPWCYTTDPEKRYDYCDILECEBCHGSGENY 193  
OY 145 --RLSTANDFNCKS-----FMYSNLTRS-CVLSDESRPLGRANLAEPV 186  
Db 194 DCKISTKWSGLEQWDSQSPHAGYIPSKRPKNLKNYCRNPDEPRP----- 243  
OY 187 GWTYFESRGVPSFTVPQMLLVGFASFMVENVPSTVMTCLDQCTSPPETGONFVC 241  
Db 244 W-----CFTTDPN-----KRMFCDIP-----RCTTPPSSSGPYQC 275

RESULT 4  
US-09-946-893-5  
; Sequence 5, Application US/09946893  
; Patent No. US20020072494A1  
; GENERAL INFORMATION:  
; APPLICANT: Cao, Yihai  
; TITLE OF INVENTION: Materials and methods relating to endothelial cell growth  
; FILE REFERENCE: Newburn  
; CURRENT APPLICATION NUMBER: US/09/946,893  
; PRIOR FILING DATE: 2001-09-05  
; PRIOR APPLICATION NUMBER: US 60/230,893  
; NUMBER OF SEQ ID NOS: 11  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 5  
; LENGTH: 569  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; OTHER INFORMATION: Description of Artificial Sequence: Angiogenesis  
US-09-946-893-5

Query Match 7.5%; Score 101; DB 9; Length 569;  
Best Local Similarity 18.0%; Pred. No. 0.16;  
Matches 53; Conservative 42; Mismatches 90; Indels 110; Gaps 13;  
OY 5 LVGFARVVSADSVHRCISACLNAPDTFGFECESVMYYPVDAECTINTEDR-----LDRP 59  
Db 33 LFSVTYKQLGAGSIECAKCEBDE---FTCRAFQYHSKEQOCVMAENKSSITIIIMR 89





```
Db 71 DVLFEKK---VYLSECKTGKNGYRGMTSKNGITQCKWSSST-----SPH----- 114
Qy 120 RQSDDELHDSFEGL-----LSCEBELCTQ----- 144
Db 115 RPRSPATHPSEGLSEENYCRNPNDPQGPWCYTTPDEKRYDCILCEEECHMGSGENY 174
Qy 145 --RLSVTANDPNCKS-----FMYSNLTRS-CVLSDESRPLGRANLAIEVP 186
Db 175 DGLSKITMSGLECCQMSQSPHAGYIPSKFPNKLKKNYCRNPRLRLP----- 224
Qy 187 GMTYFESRGVPSFTRVPQMLLVGFASFVMENVPSTVMTCLDOCTSPPTGQNFVC 241
Db 225 -W-----CFTTDPN-----KRWELCDIP-----RCTTPPSSGPTTYQC 256

RESULT 11
US-10-753-646-1
; Sequence 1, Application US/10753646
; Publication No. US20040138127A1
; GENERAL INFORMATION:
; APPLICANT: Abbott Laboratories
; TITLE OF INVENTION: NOVEL ANTIANGIOGENIC PEPTIDES,
; TITLE OF INVENTION: POLYNUCLEOTIDES ENCODING SAME AND METHODS FOR INHIBITING
; FILE REFERENCE: 5940 US P3
; CURRENT APPLICATION NUMBER: US/10/753,646
; CURRENT FILING DATE: 2004-01-08
; PRIOR APPLICATION NUMBER: US/08/924,287A
; PRIOR FILING DATE: 2004-01-08
; PRIOR APPLICATION NUMBER: US 08/851,350
; PRIOR FILING DATE: 1997-05-05
; PRIOR APPLICATION NUMBER: US 08/832,087
; PRIOR FILING DATE: 1997-04-03
; PRIOR APPLICATION NUMBER: US 08/643,219
; PRIOR FILING DATE: 1996-05-03
; NUMBER OF SEQ ID NOS: 40
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 791
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-753-646-1

Query Match 7.5%; Score 101; DB 16; Length 791;
Best Local Similarity 18.0%; Pred. No. 0.24;
Matches 53; Conservative 42; Mismatches 90; Indels 110; Gaps 13;

Qy 5 LVGFARFVNSADSVHRCISACLNAPDTFGFECESVMYYPVDAECILNTEDR-----LDRP 59
Db 14 LFSVTKKQLGAGSIEBCAKCEBDE--FTCRAPQYHSKQCCVIMENKSSIIIMR 70
Qy 60 DLFVDEHEDTVIYLDNNCAGCECHWHFDNFKTSGLINDQCPAIAQCYAPYVTVYVAVEG 119
Db 71 DVLFEKK---VYLSECKTGKNGYRGMTSKNGITQCKWSSST-----SPH----- 114
Qy 120 RQSDDELHDSFEGL-----LSCEBELCTQ----- 144
Db 115 RPRSPATHPSEGLSEENYCRNPNDPQGPWCYTTPDEKRYDCILCEEECHMGSGENY 174
Qy 145 --RLSVTANDPNCKS-----FMYSNLTRS-CVLSDESRPLGRANLAIEVP 186
Db 175 DGLSKITMSGLECCQMSQSPHAGYIPSKFPNKLKKNYCRNPRLRLP----- 224
Qy 187 GMTYFESRGVPSFTRVPQMLLVGFASFVMENVPSTVMTCLDOCTSPPTGQNFVC 241
Db 225 -W-----CFTTDPN-----KRWELCDIP-----RCTTPPSSGPTTYQC 256

RESULT 12
US-10-735-577-1
; Sequence 1, Application US/10735577
```

```
; Publication No. US20040142897A1
; GENERAL INFORMATION:
; APPLICANT: Walteman, David M.
; TITLE OF INVENTION: Compositions and Methods for Inhibiting Tumor Growth and Metastasi
; FILE REFERENCE: ME03-009
; CURRENT APPLICATION NUMBER: US/10/735,577
; CURRENT FILING DATE: 2003-12-12
; PRIOR APPLICATION NUMBER: US 60/433,140
; PRIOR FILING DATE: 2002-12-13
; NUMBER OF SEQ ID NOS: 160
; SOFTWARE: Microsoft Word
; SEQ ID NO 1
; LENGTH: 791
; TYPE: PRT
; ORGANISM: mammalian
US-10-735-577-1

Query Match 7.5%; Score 101; DB 16; Length 791;
Best Local Similarity 18.0%; Pred. No. 0.24;
Matches 53; Conservative 42; Mismatches 90; Indels 110; Gaps 13;

Qy 5 LVGFARFVNSADSVHRCISACLNAPDTFGFECESVMYYPVDAECILNTEDR-----LDRP 59
Db 14 LFSVTKKQLGAGSIEBCAKCEBDE--FTCRAPQYHSKQCCVIMENKSSIIIMR 70
Qy 60 DLFVDEHEDTVIYLDNNCAGCECHWHFDNFKTSGLINDQCPAIAQCYAPYVTVYVAVEG 119
Db 71 DVLFEKK---VYLSECKTGKNGYRGMTSKNGITQCKWSSST-----SPH----- 114
Qy 120 RQSDDELHDSFEGL-----LSCEBELCTQ----- 144
Db 115 RPRSPATHPSEGLSEENYCRNPNDPQGPWCYTTPDEKRYDCILCEEECHMGSGENY 174
Qy 145 --RLSVTANDPNCKS-----FMYSNLTRS-CVLSDESRPLGRANLAIEVP 186
Db 175 DGLSKITMSGLECCQMSQSPHAGYIPSKFPNKLKKNYCRNPRLRLP----- 224
Qy 187 GMTYFESRGVPSFTRVPQMLLVGFASFVMENVPSTVMTCLDOCTSPPTGQNFVC 241
Db 225 -W-----CFTTDPN-----KRWELCDIP-----RCTTPPSSGPTTYQC 256

RESULT 13
US-09-946-893-2
; Sequence 2, Application US/09946893
; Patent No. US20020072494A1
; GENERAL INFORMATION:
; APPLICANT: Cao, Yihai
; TITLE OF INVENTION: Materials and methods relating to endothelial cell growth
; FILE REFERENCE: Mewburn
; CURRENT APPLICATION NUMBER: US/09/946,893
; CURRENT FILING DATE: 2001-09-05
; PRIOR APPLICATION NUMBER: US 60/230,893
; PRIOR FILING DATE: 2000-09-05
; NUMBER OF SEQ ID NOS: 11
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 810
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-946-893-2

Query Match 7.5%; Score 101; DB 9; Length 810;
Best Local Similarity 18.0%; Pred. No. 0.25;
Matches 53; Conservative 42; Mismatches 90; Indels 110; Gaps 13;

Qy 5 LVGFARFVNSADSVHRCISACLNAPDTFGFECESVMYYPVDAECILNTEDR-----LDRP 59
Db 33 LFSVTKKQLGAGSIEBCAKCEBDE--FTCRAPQYHSKQCCVIMENKSSIIIMR 89
Qy 60 DLFVDEHEDTVIYLDNNCAGCECHWHFDNFKTSGLINDQCPAIAQCYAPYVTVYVAVEG 119
```

```

Db      90 DVLVFEKK--VYLSECKTGNGKNGRTGMSKTNGITCQKMSST-----SPH----- 133
Qy      120 RQLSDLDHSFEGLE-----LSECEELCTQ----- 144
Db      134 RPRFSPATHPSEGLSEENYCRNPNDPOGWCYTTDPEKRYDYCDILECEBECHGSGENY 193
Qy      145 --RLSVTANDFNCKS-----FMYSNLTRS--CVLSDESRPPLGRANLAEPV 186
Db      194 DGIKSKTWSGLECQAWDSQSPHAGYIPSKFPMKNLKKNYCRNPDRLELR----- 243
Qy      187 GWTYFESRGVPSFTTRVPQMLLVGFASFVMEVNSVTMCLDOCTSPPEPTGQNFVC 241
Db      244 -W-----CFTTDPN-----KRWELCDIP-----RCTTPPSSGPTYQC 275

```

## RESULT 14

```

US-10-193-656-2
/ Sequence 2, Application US/10193656
/ Publication No. US20030096733A1
/ GENERAL INFORMATION:
/ APPLICANT: NY, Tor
/ APPLICANT: HOLMDAHL, Richard
/ APPLICANT: LI, Jinan
/ TITLE OF INVENTION: NOVEL DRUG TARGETS FOR ARTHRITIS
/ FILE REFERENCE: 3810/10577-US3
/ CURRENT APPLICATION NUMBER: US/10/193,656
/ PRIOR FILING DATE: 2002-07-10
/ PRIOR APPLICATION NUMBER: US 60/304,461
/ PRIOR FILING DATE: 2001-07-10
/ PRIOR APPLICATION NUMBER: US 60/304,490
/ PRIOR FILING DATE: 2001-07-10
/ PRIOR APPLICATION NUMBER: US 60/305,182
/ PRIOR FILING DATE: 2001-07-13
/ NUMBER OF SEQ ID NOS: 18
/ SOFTWARE: PatentIn version 3.1
/ SEQ ID NO 2
/ LENGTH: 810
/ TYPE: PRT
/ ORGANISM: Homo sapiens
/ PUBLICATION INFORMATION:
/ DATABASE ACCESSION NUMBER: GenBank / P00747
/ DATABASE ENTRY DATE: 1986-07-21
/ RELEVANT RESIDUES: (1)..(810)
US-10-193-656-2

```

```

Query Match      7.5%; Score 101; DB 14; Length 810;
Best Local Similarity 18.0%; Pred. No. 0.25; Indels 110; Gaps 13;
Matches 53; Conservative 42; Mismatches 90;

```

```

Qy      5 LVGFAREVVSADSVHRCLSACLNAPDTFGFECESVMYYPVDAECILNTEDR-----LDRP 59
Db      33 LFSVTKKQAGASIECAKCEDEE---FTCRAPQYHSKEQOCVIMAEKRKSSIIIRMR 89
Qy      60 DLFDVDEHEDTVIYLDNNCAGCECHWFNDNFKTSIGILNDQOPALAAQCYAPVYQYVAVEG 119
Db      90 DVLVFEKK--VYLSECKTGNGKNGRTGMSKTNGITCQKMSST-----SPH----- 133
Qy      120 RQLSDLDHSFEGLE-----LSECEELCTQ----- 144
Db      134 RPRFSPATHPSEGLSEENYCRNPNDPOGWCYTTDPEKRYDYCDILECEBECHGSGENY 193
Qy      145 --RLSVTANDFNCKS-----FMYSNLTRS--CVLSDESRPPLGRANLAEPV 186
Db      194 DGIKSKTWSGLECQAWDSQSPHAGYIPSKFPMKNLKKNYCRNPDRLELR----- 243
Qy      187 GWTYFESRGVPSFTTRVPQMLLVGFASFVMEVNSVTMCLDOCTSPPEPTGQNFVC 241
Db      244 -W-----CFTTDPN-----KRWELCDIP-----RCTTPPSSGPTYQC 275

```

## RESULT 15

```

US-10-237-144-1
/ Sequence 1, Application US/10237144

```

```

/ Publication No. US20030147879A1
/ GENERAL INFORMATION:
/ APPLICANT: NY, Tor
/ APPLICANT: LI, Jinan
/ APPLICANT: HELSTROM, Sten
/ APPLICANT: ERIKSSON, Per-Olof
/ TITLE OF INVENTION: METHOD OF WOUND HEALING
/ FILE REFERENCE: 3810/10759-US1
/ CURRENT APPLICATION NUMBER: US/10/237,144
/ PRIOR FILING DATE: 2002-09-05
/ PRIOR APPLICATION NUMBER: US 60/317,643
/ PRIOR FILING DATE: 2001-09-06
/ NUMBER OF SEQ ID NOS: 5
/ SOFTWARE: PatentIn version 3.1
/ SEQ ID NO 1
/ LENGTH: 810
/ TYPE: PRT
/ ORGANISM: Homo sapiens
/ PUBLICATION INFORMATION:
/ DATABASE ACCESSION NUMBER: GenBank / 625234
/ DATABASE ENTRY DATE: 2000-09-15
/ RELEVANT RESIDUES: (1)..(810)
US-10-237-144-1

```

```

Query Match      7.5%; Score 101; DB 14; Length 810;
Best Local Similarity 18.0%; Pred. No. 0.25; Indels 110; Gaps 13;
Matches 53; Conservative 42; Mismatches 90;

```

```

Qy      5 LVGFAREVVSADSVHRCLSACLNAPDTFGFECESVMYYPVDAECILNTEDR-----LDRP 59
Db      33 LFSVTKKQAGASIECAKCEDEE---FTCRAPQYHSKEQOCVIMAEKRKSSIIIRMR 89
Qy      60 DLFDVDEHEDTVIYLDNNCAGCECHWFNDNFKTSIGILNDQOPALAAQCYAPVYQYVAVEG 119
Db      90 DVLVFEKK--VYLSECKTGNGKNGRTGMSKTNGITCQKMSST-----SPH----- 133
Qy      120 RQLSDLDHSFEGLE-----LSECEELCTQ----- 144
Db      134 RPRFSPATHPSEGLSEENYCRNPNDPOGWCYTTDPEKRYDYCDILECEBECHGSGENY 193
Qy      145 --RLSVTANDFNCKS-----FMYSNLTRS--CVLSDESRPPLGRANLAEPV 186
Db      194 DGIKSKTWSGLECQAWDSQSPHAGYIPSKFPMKNLKKNYCRNPDRLELR----- 243
Qy      187 GWTYFESRGVPSFTTRVPQMLLVGFASFVMEVNSVTMCLDOCTSPPEPTGQNFVC 241
Db      244 -W-----CFTTDPN-----KRWELCDIP-----RCTTPPSSGPTYQC 275

```

```

Search completed: March 31, 2005, 02:10:10
Job time : 142 secs

```





R:Matthews, P.  
 submitted to the EMBL Data Library, January 1996  
 A:Reference number: Z19495  
 A:Accession: T21967  
 A:Status: preliminary; translated from GB/EMBL/DBD  
 A:Molecule type: DNA  
 A:Residues: 1-751 <MTL>  
 A:Cross-references: UNIPROT:Q20167; EMBL:Z68342; PIDN:CAA92773.1; GSPDB:GN00022; CESP:F3  
 A:Experimental source: clone F38B11  
 C:Genetics:  
 A:Gene: CESP:F38B11.4  
 A:Map position: 4  
 A:introns: 23/1; 68/3; 116/1; 210/1; 284/1; 307/1; 359/3; 539/3; 622/2; 660/1; 695/2

Query Match 19.4%; Score 262; DB 2; Length 751;  
 Best Local Similarity 27.4%; Pred. No. 5.7e-16;

Matches 73; Conservative 37; Mismatches 98; Indels 58; Gaps 8;

QY 2 QKLVGFARVVSADSVHRCLSACLNAPDTFGECESVMYYVDAECIINTEDRLDRPD 61  
 DB 131 QSLIGHAMKAVLVVDGLSDCLSKALSORSYDFLCKSAIYYETGECIMNRDKFIYPL 190  
 QY 62 FVDEHEDTVI-YLDNNCAGCEC-----HHMFNFKTSGILNDQQAIAAQCYAVVTOY 114  
 DB 191 FKTNIIDTLVDYFENNCAADVSCKPEETLHW----- 220  
 QY 115 VAVGKQLSDDELHSECELETCQRLSTVANDFNCKSFMYSLNTRSCVLSDEBS 174  
 DB 221 VRTREKVLDESQVIVYESSDAQECNQLC-QNNKIGENFPCCAFASNSKQEBHLTAESS 279  
 QY 175 ----RFLGRANLAEPGWTFFESRGVP-----SFTRVPMQLVGFASFVWENVPST 222  
 DB 280 YVGHKDDKCFNMLAPLNSGEYFEKCYCLPTMLQCIASFELVANNMT--SAYKTIASLSQH 337  
 QY 223 MCLDQCTSPPEPTGQNFVCKSVWYYY 248  
 DB 338 ECLSDQCMNDGAR-----CSSATTFY 357

## RESULT 3

T22486  
 hypothetical protein F52B11.3 - Caenorhabditis elegans  
 C:Species: Caenorhabditis elegans  
 C>Date: 15-Oct-1999 #sequence\_revision 15-Oct-1999 #text\_change 09-Jul-2004  
 C:Accession: T22486  
 R:Matthews, L.  
 submitted to the EMBL Data Library, November 1996  
 A:Reference number: Z19569  
 A:Accession: T22486  
 A:Status: preliminary; translated from GB/EMBL/DBD  
 A:Molecule type: DNA  
 A:Residues: 1-741 <MTL>  
 A:Cross-references: UNIPROT:Q9XU85; EMBL:Z68266; PIDN:CAM05199.1; GSPDB:GN00022; CESP:F5  
 A:Experimental source: clone F52B11  
 C:Genetics:  
 A:Gene: CESP:F52B11.3  
 A:Map position: 4  
 A:introns: 48/1; 298/1; 341/3; 690/2

Query Match 12.4%; Score 168; DB 2; Length 741;  
 Best Local Similarity 22.5%; Pred. No. 2.5e-07;

Matches 76; Conservative 38; Mismatches 100; Indels 124; Gaps 15;

QY 5 LVGFARVVSADSVHRCLSACLNAPDTFGECESVMYYVDAECIINTEDRLDRPD 64  
 DB 137 LEGFVKKSTVENREHCLSAKKE--FVCKSVFHDITSLCELSVEDKSKRP----- 188  
 QY 65 EH---EDTVIYLDNNCAG-----C--EC 82  
 DB 189 TVRMSSEKIDYDNNCLSRQRCGPGGNLVFKTTTFIRYRHTQSVGAQESCLQKC 248  
 QY 83 HHMFNFKTSGILN-----DQO-----FALAAQCYAVVTOY 114

DB 249 LDIINTFCGSVFNENKENCIVSDEDTFSRADQOQGVGKDYEPICVAADLSSTCRQ 308  
 QY 115 VAVE---GRQSDDELHSECELETCQRLSTVANDFNCKSFMYSLNTRSCVLS 171  
 DB 309 AAFERFISGSLGEEVAVAQGVLTISDCLSCRONL-----NCKSINYDRTASSCFI-- 359  
 QY 172 ERSRPLGR--ANLAEPGWTFFESRGVPSF-----TRVPMQLVGF-- 211  
 DB 360 ---YAVGQDANIKANPMSDYEFNCEGQFGSMALCTNEGIRFIYNTKEPTGALYAAER 416  
 QY 212 ----SFMENVPSTMLDQCTSPPEPTGQNFVCKSV 245  
 DB 417 FSTGQVVENAKQISI-----TFPPTVSSD--CGTVI 447

## RESULT 4

GB9459  
 protein C52B11.1 [imported] - Caenorhabditis elegans  
 C:Species: Caenorhabditis elegans  
 C>Date: 10-May-2001 #sequence\_revision 10-May-2001 #text\_change 09-Nov-2001  
 C:Accession: GB9459  
 R:Anonymous, The C. elegans Sequencing Consortium.  
 Science 282, 2012-2018, 1998  
 A:Title: Genome sequence of the nematode C. elegans: a platform for investigating biology  
 A:Reference number: A75000; MUID:99069613; PMID:9851916  
 A>Note: see websites genome.wustl.edu/gsc/C\_elegans/ and www.sanger.ac.uk/projects/C\_elec  
 A:Note: published errata appeared in Science 283, 35, 1999; Science 283, 2103, 1999; and  
 A:Accession: GB9459  
 A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-282 <STO>  
 A:Cross-references: chr\_X; PIDN:AAA82467.1; PID:91086865; GSPDB:GN00028; CESP:C52B11.1  
 C:Genetics:  
 A:Gene: C52B11.1  
 A:Map position: X

Query Match 11.8%; Score 160; DB 2; Length 282;  
 Best Local Similarity 28.9%; Pred. No. 4.1e-07;

Matches 43; Conservative 17; Mismatches 71; Indels 18; Gaps 4;

QY 112 TOYVAVGKQLSDDELHSECELETCQRLSTVANDFNCKSFMYSLNTRSCVLS 171  
 DB 31 THFVTDNASLQSDADPIYKATSEECISACTKRDREIVCHSFYDHASFSCITLHK 90  
 QY 172 ERSRPLGRANLAEPGWTFFE-----SRGV-----SFTRVPMQLVGFASFVWENVPST 221  
 DB 91 EKSAPVGAQIENSVGKRYFEKICLSHRIPQGCQATQFIRVDQSVLVGYA-VNMTLTDST 149  
 QY 222 TMLDQCTSPPEPTGQNFVCKSVWYYYNE 250  
 DB 150 ESCAAQC-----QADCKGSANVYFED 171

## RESULT 5

T32444  
 hypothetical protein H42K12.3 - Caenorhabditis elegans  
 C:Species: Caenorhabditis elegans  
 C>Date: 29-Oct-1999 #sequence\_revision 29-Oct-1999 #text\_change 09-Jul-2004  
 C:Accession: T32444; T15821  
 R:Maggil, L., Harper, M.  
 submitted to the EMBL Data Library, September 1997  
 A:Description: The sequence of C. elegans cosmid H42K12.  
 A:Reference number: Z21169  
 A:Accession: T32444  
 A:Status: preliminary; translated from GB/EMBL/DBD  
 A:Molecule type: DNA  
 A:Residues: 1-522 <MAG>  
 A:Cross-references: UNIPROT:O17347; EMBL:AF026207; PIDN:AAB71266.1; GSPDB:GN00028; CESP:I  
 A:Experimental source: strain Bristol N2; clone H42K12  
 R:Martin, J.  
 submitted to the EMBL Data Library, November 1995  
 A:Description: The sequence of C. elegans cosmid C52B11.  
 A:Reference number: Z18411

```

A:Accession: J15821
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Reidues: 1-280, 'GK' <MAR>
A:Cross-references: EMBL:U41276; NID:g1066884; PID:g1086885; PIDN:AAA82467.1; CESP:CS2B
A:Experimental source: strain Bristol N2
C:Genetics:
A:Gene: CESP:H42K12.3; CESP:CS2B11.1
A:Map position: X
A:Introns: 24/1, 65/3, 117/1, 163/3, 208/1, 281/1, 381/3, 430/1, 481/3

Query Match      11.8%; Score 160; DB 2; Length 522;
Best Local Similarity 28.9%; Pred. No. 8.7e-07;
Matches 43; Conservative 17; Mismatches 71; Indels 18; Gaps 4;

QY      112 TQYAVNEGROSLDELHDSFEGLSECEELCTQRLSTANDFNCKSFMYSLNTRSCVLSL 171
          : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB      31 THFVLTNDASIQSDADPIYVKATSEEBCLACTKNRDKDREPIVCHSFYDHASFSCTIHK 90
          : : : : : : : : : : : : : : : : : : : : : : : : : : : :

QY*      172 ERSRPLGRANLAIEVPGWTFE-----SRGV-----SFTRPQMLIVGCAFVWENVPV 221
          : : : : : : : : : : : : : : : : : : : : : : : : : : : :
          91 EKSAIPVSAQIENSVGKRYFEKICLSHIIPOCAQOQIFIRDOSVLVGYA-VNNTLTDSI 149
          : : : : : : : : : : : : : : : : : : : : : : : : : : : :

QY      222 TMCIDQCTSPRPETGQNFVCKSWMYVNE 250
          : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB      150 ESCAQC-----QEADCKSAMFYED 171
          : : : : : : : : : : : : : : : : : : : : : : : : : : : :

```

```

RESULT 6
T29557
hypotheoretical protein C16D9.1 - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 09-Jul-2004
C:Accession: T29557
R:Gattung, S.; Le, T.T.
submitted to the EMBL Data Library, July 1996
A:Description: The sequence of C. elegans coeimid C16D9.
A:Reference number: Z20640
A:Accession: T29557
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-342 <GAT>
A:Cross-references: UNIPROT:Q22902; EMBL:U64858; PIDD:AA018288.1; GSPDB:IGN00023; CESP:CA
A:Experimental source: strain Bristol N2; clone C16D9
C:Genetics:
A:Gene: CESP:C16D9.1
A:Map position: 5
A:introns: 59/2, 316/3

Query Match      11.3%; Score 153.5; DB 2; Length 342;
      Basic local Similarity 38.2%; Pred. No. 2.1e-06;
      Matches 29; Conservative 18; Mismatches 26; Indels 3; Gaps 2;

OY      3 KILVFAAEVVA-SADSVARCLSAACINADPTGFECESTWYVPVPAECIIINTEDRLDRPDL 61
      Db -      121 KVLIGIVDQLRVDAIVAVDCAQCSNSQTXYDITCKSAMYYEKQKECIIASQSRADIPDL 180
      OY      62 FVDEHEEDVITYLDNNC 77
      Db      181 FID--DDKSLYLENSC 194

RESULT 7
T23990
hypotheoretical protein R07A4.4 - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 09-Jul-2004
C:Accession: T23990
R:Cottage, A.
submitted to the EMBL Data Library, November 1995
A:Reference number: Z19827
A:Accession: T23990
A:Status: preliminary; translated from GB/EMBL/DBJ

```

A:Molecule type: DNA  
A:Residues: 1-351 <WIL>  
A:Cross-references: UNIPROT:Q21782, EMBL:Z6756, PDB:CAA91764.1, GSPDB:GN00028, CESP:R07A4  
A:Experimental source: clone R07A4  
C:Genetics:  
A:Gene: CESP:R07A4.4  
A:Map position: X  
A:introns: 24/3, 47/3, 72/3, 126/1, 147/1, 169/3, 227/1, 247/1, 293/3, 321/1

Query Match	9.3%	Score 126;	DB 2;	Length 351;
Best Local Similarity	25.0%	Pred. No. 0.00072;		
Matches	45;	Conservative	31;	Mismatches 72; Indels 32; Gaps 7
QY	95	INDOO-----FA-----IAOCYAPVYQYAVBGRQSLDELHSEGLELSECEEL	141	
DB	8	INQQQDFNTLRNFAVEVLPIVPDPCPLGSHSRVQILEGVEGRBATITFQVAILLEQCVQA	67	
QY	142	CTQRLSTVANDFN---CKSFMYISLNTLRSCTLSDRSRPLGRANLAEPGMYTFESRGVP	197	
DB	68	C--RVSTYADSDRLPLCLCRSAHFNFNATRCQSYSDAIPNGLYKXPNOMYIIEKICIP	125	
QY	198	S-----FTFVPMQLVGFSPFVAVENVPSTMCDDOCTSPPEPTGONFVCKSVMY	247	
DB	126	DTVLPMGCDVFRKRPQHLGHANSEV--SVASEECVLECK--AKTLRSVACHSILHY	182	

RESULT 8  
T25804  
hypothetical protein H03E18.1 - *Caenorhabditis elegans*  
C/Species: *Caenorhabditis elegans*  
C/Date: 15-Oct-1999 #sequence\_revision 15-Oct-1999 #text\_change 09-Jul-2004  
C/Accession: T25804  
R/Bentley, D.  
submitted to the EMBL Data Library, August 1996  
A/Description: The sequence of C. elegans cosmid H03E18.  
A/Reference number: 220090  
A/Accession: T25804  
A/Status: preliminary; translated from GB/EMBL/DBJ  
A/Molecule type: DNA  
A/Residues: 1-1127 <BEN>  
A/Cross-References: UNIPROT:Q94248; EMBL:U67947; RIND:AA07557.1; GSPDB:GN00028; CESP:H03E18  
A/Experimental source: strain Bristol NZ; clone H03E18  
C/Genetics:  
A/Gene: CESP:H03E18.1  
A/Map position: X  
A/Introns: 112/3; 217/1; 282/2; 304/1; 424/1; 461/2; 579/3; 615/2; 831/1; 999/2; 1088/1

Query Match	9.2%	Score 134.5	DB 2	Length 1127
Best Local Similarity	22.3%	Pred. No. 0.0041		
Matches	44	Conservative	37	Mismatches 75
			Indels	41
			Gaps	9

  

QY	1	EOKILVGFAREVVSADSVHRCISACLNAPDT--FGPECESVNYYPVDAECILNTEDRDLR	58
		: : : :	
DB	433	DDHLVSVSAGGLEHMDSIIECCMCANSKTSRRYFQCASATYYHAERDCILNLEDRNIK	4922
		: : : :	
QY	59	PDLFVDEHED-TVLYLDNNC-----AGECHMHNP-----NPKSGIL	95
		: : : :	
DB	493	SKLEPEKQTTDNNVYITGLACEYDETVKSLGSLATECRKRVZATYTGEPFKETSTKNGGLK	552
		: : : :	
QY	96	NDQGFALMAOCYAPVYQYVAVEGRHLSDELHDSFEGELSSCEBELCTQRLSVTANDPNC	155
		: : : :	
DB	553	SD-----DCYVE-LNDFV-LEGTAIAVE-----TAVTPRECKCKCAEGHKLGYEE--C	596
		: : : :	
QY	156	KSPMYNSNTRSCVSLDE	172
		: : : :	
DB	597	ASFLYYDSKTCCLINKQ	613
		: : : :	

RESULT 9  
PLPG  
plasmin (EC 3.4.21.7) precursor - pig (fragment  
N:Alternate names: plasminogen  
N:Contains: miniplasminogen



```

Db      94 RMRDYLVEKRIYLLECKTNGQTFRTAETKSGVTCOKMSATSPHYPKF-----SPE 147
      99 QFAIAA-----QCYAP-----YVTQYAVEGQSLDELHDSFEGLESECELTQ- 144
      148 KFLPLGLEENYCRNDNDENGPMCTT-----DPPKRYDYCDIECECEKCMHC 195
      145 -----RLSYTANDFNCKS-----FMYSNLTLS-CVLSDESRPLGRAN 181
      196 GGENVEGIATKMSGRDQAMDOSPHAHGYIPSKFPMKNLKMVCRNPDEPR----- 250
      182 LAEVPGWTFESRGVPSFTRVPQMLLVGFASFVMEVNSVTMCLDOCTSPPETGQNFVC 241
      251 -----W-----CFTTDPQ-----KRMFCDIP-----RCTTPPSSSGPKYQC 282

```

# RESULT 11

```

B30848
Plasmin (EC 3.4.21.7) precursor - rhesus macaque
C/Species: Macaca mulatta (rhesus macaque)
C/Date: 31-Mar-1989 #sequence_revision 31-Mar-1989 #text_change 09-Jul-2004
C/Accession: B32866; B30848
R/Tomlinson, J.E.; McLean, J.W.; Lawn, R.M.
J. Biol. Chem. 264, 5957-5965, 1989
A/Title: Rhesus monkey apolipoprotein(a). Sequence, evolution, and sites of synthesis.
A/Reference number: A32869; MUID:89174660; PMID:2925643
A/Accession: B32869
A/Status: preliminary
A/Molecule type: mRNA
A/Residues: 1-810 <TOM>
A/Cross-references: UNIPROT:P12545; GB:J04697; NID:9342272; PIDN:AAA36901.1; PID:9342273
C/Superfamily: Plasmin; kringle homology; plasminogen-related protein precursor homology
C/Keywords: fibrinolysis; glycoprotein; hydrolyase; kringle; serine proteinase
F:1-96/Domain: plasminogen-related protein precursor homology <PIPH>
F:1-9/Domain: signal sequence #status predicted <SIG>
F:103-181/Domain: kringle homology <KR1>
F:185-262/Domain: kringle homology <KR2>
F:275-352/Domain: kringle homology <KR3>
F:377-454/Domain: kringle homology <KR4>
F:481-560/Domain: kringle homology <KR5>
F:581-803/Domain: trypsin homology <TRY>
F:49-73,53-61,103-181,124-164,152-176,185-262,188-316,206-245,234-257,275-352,296-335,32
bonds: #status predicted
F:622,665,760/Active site: His, Asp, Ser #status predicted

```

```

Query Match      7.7%; Score 104; DB 2; Length 810;
Best Local Similarity 18.1%; Pred. No. 0.21;
Matches 54; Conservative 38; Mismatches 89; Indels 118; Gaps 14;

```

```

      5 LVGFARFVVSADSVRCISACLNADFTFGFECESVWYYPVDAECTINTEDRLDRPDLFVD 64
      33 LFSITTKQLGAGSIECAKCEBEEB---FTCRSQYHSKQOQCVIAE---NRKSIVF 86
      65 EHEBTVIYLDNNCAGCECHMFDFNFKTS-----GILNDQOFAIAAQC---APVYQYV 115
      87 RMRDVLVE-----EKVYVLSCKTNGKNGVGTMSKTTGITCOKMSSTSH----- 133
      116 AVEGRQLSDELDHSEGLE-----LSCEBELCTQ-- 144
      134 -----RPTFSPTHSEGLEENYCRNPNDGQPMCYTTDPERPDYCDIPCECEBCHMCS 189
      145 -----RLSYTANDFNCKS-----FMYSNLTLS-CVLSDESRPLGRANL 182
      190 GENVGDKSKTMSGIECCAMDSQSPHAGYIPSKFPMKNLKMVCRNPDEGRPR----- 243
      183 AEVPGWTFESRGVPSFTRVPQMLLVGFASFVMEVNSVTMCLDOCTSPPETGQNFVC 241
      244 -----W-----CFTTDPN-----KRMELCIDP-----RCTTPPSSSGPKYQC 275

```

RESULT 12  
T16417  
hypothetical protein F52C9.5 - Caenorhabditis elegans

```

C/Species: Caenorhabditis elegans
C/Date: 20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change 09-Jul-2004
C/Accession: T16417
R/Favella, T.
Submitted to the EMBL Data Library, November 1995
A/Description: The sequence of C. elegans cosmid F52C9.
A/Reference number: 218511
A/Accession: T16417
A/Status: preliminary; translated from GB/EMBL/DBJ
A/Molecule type: DNA
A/Residues: 1-498 <FAV>
A/Cross-references: UNIPROT:Q10125; EMBL:U39850; NID:g1055052; PID:g1055053; PIDN:AAA810;
C/Superfamily: Caenorhabditis elegans hypothetical protein F52C9.5
A/Introns: 45/2; 87/3; 132/1; 185/3; 207/1; 294/3; 337/1; 381/3; 430/1

```

```

Query Match      7.6%; Score 103; DB 2; Length 498;
Best Local Similarity 19.9%; Pred. No. 0.14;
Matches 53; Conservative 27; Mismatches 65; Indels 122; Gaps 11;

```

```

      14 SADSVHRCISACLNADFTFG-FECESVWYYPVDAECTL----- 50
      154 SSTGLHCKSHCLNS--QIGVSCRSFVYDVMNRVCDLFAHVGDQAPARLKFQTOEDS 211
      51 -----NTEDRLDRPDLFVDEHEDTVIYLDNNCAGCECHMFDFNFKTSGL 95
      212 PPSPPSAPVIALATNTDKR-----DEHEBTEITED----- 243
      96 NDQOFAIAA---OCYAPVYQYAVEGQSLDELHDSFEGLESECELTORLSVTAN 151
      244 ----ITVASPSDSCPRGKQSTFLRTGQFLFSHDDQELVGVDAECAKACIEN----- 293
      152 DFNKSKFMSYSLTSCVLSDESRPLGRANLAEVPGWTFESRGVPS--FTRVPQMLLVG 209
      294 ----KKSPPY-----RSC-----PSYFESRHPQMLLVG 316
      210 FASFVMEVNSVTMCLDOCTSPPETG 236
      317 FAEVSQDS-PSFEHCFTCLNKTDATG 342

```

# RESULT 13

```

T19553
hypothetical protein C29B6.4 - Caenorhabditis elegans
C/Species: Caenorhabditis elegans
C/Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 09-Jul-2004
C/Accession: T19553
R/Dobson, R.
Submitted to the EMBL Data Library, May 1996
A/Reference number: Z19141
A/Accession: T19553
A/Status: preliminary; translated from GB/EMBL/DBJ
A/Molecule type: DNA
A/Residues: 1-942 <WIL>
A/Cross-references: UNIPROT:Q18298; EMBL:Z72504; PIDN:CAA96604.1; GSPDB:GN00022; CESP:C2;
A/Experimental source: clone C29B6
C/Genetics:
A/Gene: CESP:C29B6.4
A/Map position: 4
A/Introns: 21/3; 67/3; 145/2; 175/1; 243/3; 308/1; 415/2; 540/1; 616/3; 659/3; 708/2; 83;

```

```

Query Match      7.6%; Score 102.5; DB 2; Length 942;
Best Local Similarity 22.1%; Pred. No. 0.35;
Matches 36; Conservative 30; Mismatches 72; Indels 25; Gaps 6;

```

```

      13 VSADSVHRCISACLNADFTFGFECESVWYYPVDAECTINTEDRLDRPDLFVDEHEDTVIY 72
      306 ITNOSTKCLQKCTDMEN-----CKSVYVNVLTSCILSTSKISATV-SDENYD---F 356
      73 LDNNCAGCECHMFDFNFKTSGLNDQOFAIAAQCAYPYTQYAAV-GRQLSDELDHSE 131
      357 YERSC-----PYIPTSTSSSFP1PHPTLYPEVTVRYKLMERGRQLSSFVAKTN 407

```

QY 132 GLEISCELCIORLSVTANDPCKSFMYSNLRGVLSDERS 174  
 DB 408 VENVODPCMSLCVN-----SKVHCELSIFSTSTSNOCCLSTLKS 444  
 RESULT 14  
 PLHM (EC 3.4.21.7) precursor [validated] - human  
 N/Alternate names: plasminogen precursor [msmome]  
 N/Contents: angiotatin; microplasmin; plasminogen  
 C/Species: Homo sapiens (man)  
 C/Date: 24-Apr-1984 #sequence revision 02-Dec-1994 #text change 09-Jul-2004  
 C/Accession: A35229; I52242; A26646; I62733; I84609; S03735; A00929; A04627; A04625; A04  
 R/Petersen, T.E.; Martzen, W.R.; Ichinose, A.; Davie, E.W.  
 J. Biol. Chem. 265, 6104-6111, 1990  
 A/Title: Characterization of the gene for human plasminogen, a key proenzyme in the fibr  
 A/Reference number: A35229; MUID:90202879; PMID:2318848  
 A/Accession: A35229  
 A/Molecule type: DNA  
 A/Residues: 1-810 <P>  
 A/Cross-references: UNIPROT:P00747; UNIPROT:Q9UBQ9; UNIPROT:Q9UMI2; GB:J05286; GB:M34276  
 A/Experimental source: leukocyte; lung fibroblast  
 R/Malgaroli, N.; Bruno, L.; Pontoglio, M.; Candiani, G.; Meroni, G.; Ottolenghi, S.; Ta  
 Biochem. Biophys. Res. Commun. 173, 1013-1018, 1990  
 A/Title: Definition of the transcription initiation site of human plasminogen gene in li  
 A/Reference number: I52242; MUID:91097523; PMID:2268308  
 A/Accession: I52242  
 A/Status: translated from GB/EMBL/DBJ  
 A/Molecule type: DNA  
 A/Residues: 1-16 <MAL1>  
 A/Cross-references: GB:M62890; NID:9190092; PIDN:AAA3454.1; PID:G553613  
 R/Forsgren, M.; Raden, B.; Israelsson, M.; Larsson, K.; Heden, L.O.  
 FEBS Lett. 213, 254-260, 1987  
 A/Title: Molecular cloning and characterization of a full-length cDNA clone for human pl  
 A/Reference number: A26646; MUID:87162490; PMID:3030813  
 A/Accession: A26646  
 A/Molecule type: mRNA  
 A/Residues: 1-471; 'D', 473-810 <FOR>  
 A/Cross-references: GB:X05199; NID:935530; PIDN:CAA2883.1; PID:G35531  
 A/Experimental source: liver  
 R/Malinowski, D.P.; Sadler, J.E.; Davie, E.W.  
 Biochemistry 23, 4243-4250, 1984  
 A/Title: Characterization of a complementary deoxyribonucleic acid coding for human and  
 A/Reference number: I45961; MUID:85023311; PMID:6148961  
 A/Accession: I62738  
 A/Status: translated from GB/EMBL/DBJ  
 A/Molecule type: mRNA  
 A/Residues: 292-471; 'D', 473-810 <MAL2>  
 A/Cross-references: GB:X02922; NID:9190112; PIDN:AAA60124.1; PID:G387031  
 A/Accession: I84609  
 A/Status: translated from GB/EMBL/DBJ  
 A/Molecule type: DNA  
 A/Residues: 357-419 <MAL3>  
 A/Cross-references: GB:X02921; NID:9190110; PIDN:AAA60123.1; PID:G190111  
 R/Brunsholz, R.A.; Lerch, P.G.; Schaller, J.; Rickli, E.E.; Lergler, W.; Manneberg, M.;  
 Eur. J. Biochem. 114, 465-470, 1981  
 A/Title: Comparison of the primary structure of the N-terminal CNBr fragments of human,  
 A/Reference number: S03735; MUID:81212097; PMID:7228497  
 A/Accession: S03735  
 A/Molecule type: protein  
 A/Residues: 20-71; 'E', 73-76 <BRU>  
 R/Sottrop-Jensen, L.; Petersen, T.E.; Magnusson, S.  
 submitted to the Atlas, July 1977  
 A/Reference number: A00929  
 A/Accession: A00929  
 A/Molecule type: protein  
 A/Residues: 20-71; 'E', 73-85, 87-106; 'D', 108-360; 'E', 362-810 <SOT>  
 R/Wiman, B.  
 Eur. J. Biochem. 76, 129-137, 1977  
 A/Title: Primary structure of the B-chain of human plasmin.  
 A/Reference number: A04627; MUID:77225245; PMID:1142009  
 A/Accession: A04627

A/Molecule type: protein  
 A/Residues: 581-810 <MIL>  
 R/Wiman, B.; Wallen, P.  
 Eur. J. Biochem. 50, 489-494, 1975  
 A/Title: Structural relationship between "glutamic acid" and "lysine" forms of human plas  
 A/Reference number: A04625; MUID:75093329; PMID:122932  
 A/Accession: A04625  
 A/Molecule type: protein  
 A/Residues: 20-50; 'Q', 51-71; 'E', 73-85, 87-100 <M12>  
 R/Wiman, B.; Wallen, P.  
 Eur. J. Biochem. 58, 539-547, 1975  
 A/Title: Amino-acid sequence of the cyanogen-bromide fragment from human plasminogen that  
 A/Reference number: A04626; MUID:76043692; PMID:126863  
 A/Accession: A04626  
 A/Molecule type: protein  
 A/Residues: 483-507; 'E', 509-604 <M13>  
 R/Robbins, K.C.; Bernabe, P.; Arzadon, L.; Summaria, L.  
 J. Biol. Chem. 248, 1631-1633, 1973  
 A/Title: The primary structure of human plasminogen. II. The histidine loop of human plas  
 A/Reference number: A92125; MUID:73149248; PMID:4694729  
 A/Contents: annotation; active site  
 R/Groekopf, W.R.; Summaria, L.; Robbins, K.C.  
 J. Biol. Chem. 244, 3590-3597, 1969  
 A/Title: Studies on the active center of human plasmin. Partial amino acid sequence of a  
 A/Reference number: A92048; MUID:69234739; PMID:4240117  
 A/Contents: annotation; active site  
 R/Trexler, M.; Vail, Z.; Patchy, L.  
 J. Biol. Chem. 257, 7401-7406, 1982  
 A/Title: Structure of the omega-aminocarboxylic acid binding sites of human plasminogen.  
 A/Reference number: A92382; MUID:82213905; PMID:6919539  
 A/Contents: annotation; omega-aminocarboxylic acid binding sites  
 J/Vail, Z.; Patchy, L.  
 J. Biol. Chem. 259, 13690-13694, 1984  
 A/Title: The fibrin-binding site of human plasminogen. Arginines 32 and 34 are essential  
 A/Reference number: A92458; MUID:85054794; PMID:6094526  
 A/Contents: annotation; fibrin binding site; omega-aminocarboxylic acid binding site  
 J/Cao, Y.; Ji, R.W.; Davidson, D.; Schaller, J.; Marti, D.; Seehndel, S.; McCance, S.G.;  
 J. Biol. Chem. 271, 29464-29467, 1996  
 A/Title: Kringla domains of human angiotatin. Characterization of the anti-proliferative  
 A/Reference number: A58811; MUID:97067211; PMID:8910613  
 A/Contents: annotation  
 R/Ulljen, H.R.; Uguw, F.; Bini, A.; Collen, D.  
 Biochemistry 37, 4699-4702, 1998  
 A/Title: Generation of an angiotatin-like fragment from plasminogen by stromelysin-1 (M  
 A/Reference number: A58812; MUID:9548733; PMID:9548733  
 A/Contents: annotation  
 R/Tulinsky, A.; Mulichak, A.M.  
 submitted to the Brookhaven Protein Data Bank, July 1991  
 A/Reference number: A51341; PDB:1PK4  
 A/Contents: annotation; X-ray crystallography, 1.9 angstroms, residues 376-454  
 R/Tulinsky, A.; Wu, T.P.  
 submitted to the Brookhaven Protein Data Bank, July 1991  
 A/Reference number: A51488; PDB:2PK4  
 A/Contents: annotation; X-ray crystallography, 2.25 angstroms, residues 375-454  
 R/Wu, T.P.; Tulinsky, A.  
 submitted to the Brookhaven Protein Data Bank, August 1993  
 A/Reference number: A51911; PDB:1PKR  
 A/Contents: annotation; X-ray crystallography, 2.48 angstroms, residues 102-181  
 R/Padmanabhan, K.; Tulinsky, A.  
 submitted to the Brookhaven Protein Data Bank, April 1994  
 A/Reference number: A52408; PDB:1PKK  
 A/Contents: annotation; X-ray crystallography, 2.25 angstroms, residues 377-454  
 R/Tulinsky, A.; Mathews, I.I.  
 submitted to the Brookhaven Protein Data Bank, December 1995  
 A/Reference number: A65244; PDB:1CEA  
 A/Contents: annotation; X-ray crystallography, 2.1 angstroms, residues 102-181  
 R/Tulinsky, A.; Mathews, I.I.  
 submitted to the Brookhaven Protein Data Bank, December 1995  
 A/Reference number: A65245; PDB:1CEB  
 A/Contents: annotation; X-ray crystallography, 2.1 angstroms, residues 102-181  
 R/Mulichak, A.M.; Tulinsky, A.; Ravichandran, K.G.  
 Biochemistry 30, 10576-10586, 1991

A:Title: Crystal and molecular structure of human plasminogen kringle 4 refined at 1.9 Å  
A:Reference number: A58819; MUID:92031502; PMID:1657148  
A:Contents: annotation  
R:Mu, T.P.; Padmanabhan, K.; Tulinsky, A.; Mullick, A.M.  
Biochemistry 30, 10589-10594, 1991  
A:Title: The refined structure of the epsilon-aminocaproic acid complex of human plasmin  
A:Reference number: A58818; MUID:92031503; PMID:1657149  
A:Contents: annotation  
R:de Vos, A.M.; Ultsch, M.H.; Kelley, R.F.; Padmanabhan, K.; Tulinsky, A.; Westbrook, M.  
Biochemistry 31, 2702-279, 1992  
A:Title: Crystal structure of the kringle 2 domain of tissue plasminogen activator at 2.0 Å  
A:Reference number: A39483; MUID:92118803; PMID:1310033  
A:Contents: annotation; X-ray crystallography, 2.4 angstroms  
R:Stec, B.; Teeter, M.M.; Whitlow, M.; Yamano, A.  
Submitted to the Brookhaven Protein Data Bank, June 1995  
A:Reference number: A65980; PDB:1KXN  
A:Contents: annotation; X-ray crystallography, 1.67 angstroms, residues 376-454  
R:Rejzante, M.; Llinas, M.  
Submitted to the Brookhaven Protein Data Bank, August 1996  
A:Reference number: A65803; PDB:1HPJ  
A:Contents: annotation; conformation by (1)H-NMR, residues 103-181  
R:Rejzante, M.; Llinas, M.  
Submitted to the Brookhaven Protein Data Bank, August 1996  
A:Reference number: A65804; PDB:1HPK  
A:Contents: annotation; conformation by (1)H-NMR, residues 103-181  
R:Rejzante, M.R.; Llinas, M.  
Eur. J. Biochem. 221, 927-937, 1994  
A:Title: (1)H-NMR assignments and secondary structure of human plasminogen kringle 1.  
A:Reference number: S43645; MUID:94237157; PMID:8181475  
A:Contents: annotation; conformation by (1)H-NMR, residues 96-184  
R:Rejzante, M.R.; Llinas, M.  
Eur. J. Biochem. 221, 939-949, 1994  
A:Title: Solution structure of the epsilon-aminohexanoic acid complex of human plasminogen kringle 1.  
A:Reference number: A58817; MUID:94237158; PMID:8181476  
A:Contents: annotation; conformation by (1)H-NMR  
C:Comment: Plasminogen is synthesized by the kidney and is present in plasma and many other tissues.  
C:Comment: Plasminogen is converted to plasmin by plasminogen activators (see PIR:UKHU a d PIR:FGHUGB).  
C:Comment: Plasmin is inactivated by alpha-2-antiplasmin (see PIR:ITHUA2) immediately after release, resulting in two chains connected by two disulfide bonds. Without the inhibitor, C:Comment: Microplasmin is formed by autocatalytic cleavage of plasmin under artificial conditions.  
C:Comment: Streptolysin 1 (see PIR:KCHUS1) acts on plasminogen to produce angiotensin. Td cing solid tumors.  
C:Genetics:  
A:Gene: GDB:PLG  
A:Cross-references: GDB:119498; OMIM:173350  
A:Map position: 6q26-6q27  
A:Introns: 17/1; 62/2; 98/1; 136/2; 183/1; 223/2; 263/1; 317/2; 366/1; 419/2; 480/1; 529  
C:Function:  
A:Description: dissolves the fibrin of blood clots; acts as a proteolytic factor in a variety of the walls of the graafian follicle; also activates the urokinase-type plasminogen activator.  
A:Pathway: fibrinolysis  
C:Superfamily: plasmin; kringle homology; plasminogen-related protein precursor homology  
C:Keywords: angiogenesis inhibitor; blood; duplication; fibrinolysis; glycoprotein; hydrolysis  
F:1-96/Domain: plasminogen-related protein precursor homology <PLPH>  
F:1-19/Domain: signal sequence #status predicted <SIG>  
F:20-810/Product: plasminogen #status experimental <PRO>  
F:20-96/Domain: activation peptide #status experimental <APT>  
F:79-466/Product: angiotensin #status experimental <AST>  
F:97-580,581-810/Product: plasmin #status experimental <MAT>  
F:97-580/Domain: plasmin chain A #status experimental <CHA>  
F:103-181/Domain: kringle homology <KR1>  
F:185-262/Domain: kringle homology <KR2>  
F:275-352/Domain: kringle homology <KR3>  
F:377-454/Domain: kringle homology <KR4>  
F:481-560/Domain: kringle homology <KR5>  
F:550-580,581-810/Product: microplasmin #status experimental <MMT>  
Query Match 7.5%; Score 101; DB 1; Length 810;  
Best Local Similarity 18.0%; Pred. No. 0.4;  
Matches 53; Conservative 42; Mismatches 90; Indels 110; Gaps 13;

5 LVGFARVVVADSVHRCISACINADPTGFGFEGESVVMYYPVDAECILNTEDR-----IDRP 59

Db 33 LFSVTKKQLGAGSIEECAAKEEDEE---FTCRFAQYHKSKEQCIVIMAEKSSIIIRMR 89  
Qy 60 DLPLDEHEDVYIYIDNNCAGCECHWHPDNFTSGILNDQCALAQAACAPVYTOYVAVG 119  
Db 90 DVVLFEKK---VYLSECKTGKKNVGRGTMSTKXGICQKWSST---SPH----- 133  
Qy 120 RQLDELDSPEGLE-----LSECEELCTQ----- 144  
Db 134 RPRSPATHPSGELEENYCRNPNDPOGPMCTTDPPEKRYCYILCEEECHMGSGENY 193  
Qy 145 --RLSVTANDFNCKS-----FMYSNLTRS--CVLSDESRPLGRANLAEVP 186  
Db 194 DGIKSKTMGSLGECAMDSQSPHAGYIISKRPNNLKKVYCRNDRDLRP----- 243  
Qy 187 GWTFESRGVSPFRVQMLVGFASVMEVSVTMCDOCTSPPTPTGQNPVC 241  
Db 244 -W-----CFYTDPN-----KRWELCDIP-----RCTTPPPSSGPTYOC 275  
RESULT 15  
1725288  
hypothetical protein T26C5.2 - Caenorhabditis elegans  
C:Species: Caenorhabditis elegans  
C:Date: 15-Oct-1999 #sequence\_revision 15-Oct-1999 #text\_change 09-Jul-2004  
C:Accession: T25288  
R:Thomas, K.  
Submitted to the EMBL Data Library, August 1995  
A:Reference number: Z20011  
A:Accession: T25288  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: DNA  
A:Residues: 1-554 <NLS>  
A:Cross-references: UNIPROT:Q22815; EMBL:Z50859; PIDD:CAA90727.1; GSPDB:GN00020; CESP:T24  
A:Experimental source: clone T26C5  
C:Genetics:  
A:Gene: CESP:T26C5.2  
A:Map position: 2  
A:Introns: 20/1; 59/3; 103/1; 266/1; 333/1; 376/3; 542/2  
C:Superfamily: Caenorhabditis elegans hypothetical protein T26C5.2  
Query Match 7.3%; Score 99.5; DB 2; Length 554;  
Best Local Similarity 28.8%; Pred. No. 0.35;  
Matches 19; Conservative 11; Mismatches 33; Indels 3; Gaps 1;  
Qy 16 DSVHRCISACINADPTGFGFEGESVVMYYPVDAECILNTEDRLRDLPLVDEHEDVYIYIDN 75  
Db 456 DGMQLCIELCVLSTK---FTCRSSTPNPITGQCRILMTEDSWTSPDSFEYDFQKALYFEN 512  
Qy 76 NCAGE 81  
Db 513 GCTNAE 518  
Search completed: March 31, 2005, 02:06:55  
Job time : 45 secs



This Page Blank (uspto)



GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 31, 2005, 01:56:09 ; Search time 177 Seconds  
(without alignments)  
723.276 Million cell updates/sec

Title: US-10-771-708-11  
Perfect score: 1354  
Sequence: 1 EQKLIVGFARVVSADSVHR.....PPETGQNPVCKSVMYYYNE 250

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 1612378 seqs, 512079187 residues  
Total number of hits satisfying chosen parameters: 1612378

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : UniProt\_03:\*  
1: uniprot\_sprot:\*  
2: uniprot\_tramb:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	912.5	67.4	1065	2	001494
2	912.5	67.4	1069	2	083387
3	262	19.4	751	2	020167
4	168	12.4	741	2	09XUES
5	160	11.8	522	2	017347
6	157	11.6	1494	2	07PTE3
7	153.5	11.3	398	2	022902
8	138	10.2	1549	2	08MKZ3
9	138	10.2	1549	2	09BMD4
10	138	10.2	1557	2	09BMD5
11	138	10.2	1557	2	09V5X3
12	126	9.3	351	2	021782
13	125	9.2	774	2	09V9X1
14	124.5	9.2	1127	2	094248
15	121	8.9	744	2	09VAG2
16	120.5	8.9	432	2	07PLU3
17	120.5	8.9	601	2	09W520
18	120.5	8.9	715	2	08MS37
19	119.5	8.8	633	2	07OBU4
20	118.5	8.8	695	2	07ODP6
21	116	8.6	790	1	PLMN_PIG
22	115	8.5	805	2	081GS2
23	113	8.3	833	3	09V9X0
24	108	8.0	812	1	PLMN_RAT
25	105	7.8	812	1	PLMN_BOVIN
26	104	7.7	810	1	PLMN_MACMU
27	103	7.6	498	1	YSMS_CABEL
28	102.5	7.6	942	2	018298
29	101	7.5	810	1	PLMN_HUMAN
30	99.5	7.3	524	2	022815
31	99	7.3	466	2	06RC10

32	99	7.3	510	2	060123	060123 echizocacch
33	99	7.3	693	2	027394	027394 caenorhabdi
34	99	7.3	812	1	PLMN_MOUSE	P20918 mus musculu
35	98	7.2	1580	2	07RME4	07RME4 plasmodium
36	96.5	7.1	806	1	PLMN_MACEU	018783 macropus eu
37	96	7.1	627	2	064713	064713 arabidopsis
38	96	7.1	627	2	064716	064716 arabidopsis
39	96	7.1	627	2	084LS0	084LS0 arabidopsis
40	96	7.1	824	2	066S04	066S04 oikopleura
41	95.5	7.1	696	2	061814	061814 caenorhabdi
42	94.5	7.0	1615	2	07QZ09	07QZ09 giardia lam
43	94	6.9	429	2	072206	072206 bacillus ce
44	94	6.9	786	2	062201	062201 caenorhabdi
45	93.5	6.9	480	2	06G2L3	06G2L3 bartonella

ALIGNMENTS

RESULT 1  
001494 PRELIMINARY; PRT: 1065 AA.  
ID 001494  
AC 001494;  
DT 01-JUL-1997 (TREMBLrel. 04, Created)  
DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)  
DT 01-OCT-2003 (TREMBLrel. 25, Last annotation update)  
DE Hypothetical protein C3466.6.  
GN Name=C3466.6; ORFNames=C3466.6;  
OS Caenorhabditis elegans.  
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;  
OC Rhabditidae; Pelodierinae; Caenorhabditis.  
NC NCB1\_TaxID=6239;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=Bristol N2;  
RX MEDLINE=99069613; PubMed=9851916;  
RG WormBase Consortium;  
RT "Genome sequence of the nematode C. elegans: a platform for investigating biology. The C. elegans Sequencing Consortium.,"  
RL Science 282:2012-2018(1998).  
RN [2]  
RP SEQUENCE FROM N.A.  
RC STRAIN=Bristol N2;  
RA Langston Y., Rohlfing T.;  
RT "The sequence of C. elegans cosmid C3466.,"  
RL Submitted (APR-1997) to the EMBL/GenBank/DBJ databases.  
RN [3]  
RP SEQUENCE FROM N.A.  
RC STRAIN=Bristol N2;  
RA Waterston R.;  
RL Submitted (DEC-2002) to the EMBL/GenBank/DBJ databases.  
RN [4]  
RP SEQUENCE FROM N.A.  
RC STRAIN=Bristol N2;  
RG WormBase Consortium;  
RL Submitted (SRP-2004) to the EMBL/GenBank/DBJ databases.  
DR EMBL: U97407; AAB52479.2; -.  
DR PIR: C87789; C87789.  
DR WormBase; WBGene0016422; C3466.6.  
DR WormPep; C3466.6a; CE29699.  
DR InterPro; IPR001507; Endoglin/CD105.  
DR InterPro; IPR003014; PAN.  
DR InterPro; IPR003609; Pan\_app.  
DR Pfam; PF00024; PAN; 6.  
DR Pfam; PF00100; Zona\_pellucida; 1.  
DR SMART; SM00473; PAN\_AP; 6.  
DR SMART; SM00241; ZP; 1.  
DR PROSITE; PS50948; PAN; 6.  
KW Hypothetical protein.  
SQ SEQUENCE 1065 AA; 117427 MW; DB8597B650648F84 CRC64;  
Query Match 67.4%; Score 912.5; DB 2; Length 1065;  
Best Local Similarity 66.2%; Pred. No. 8.6e-76;

	Matches	172;	Conservative	33;	Mismatches	36;	Indels	19;	Gaps	5
QY	1	EOKLIVGARAREVNSADSVHRCLSACTINAFDTRFGFECESVMYTPVDAECLINTEDRLRPD	60							
Db	114	EQRLVIGARAREVVPANPVIQICMAACLAFLDTFGFECESAMFPVQECILNTEDELDLPS	173							
QY	61	LFVDEHEPTVLYLDDNKGAGCECHMFDFNFKTSGILNDOOPALA-AQCAPVYOVVAEG	119							
Db	174	LFVESDDTVLYMNNKNGAFPL--VFKEY-----NYQKTFPSQCTPPYITQIIVAE	225							
QY	120	ROLSDELHSEFGELESCBELCTQRLSVTANDFNCKSPMYSNLTRSCVLSDERSRPIGR	179							
Db	226	KQLNNELDRII-NYDLDSQALCTQRLSISNDFCKSPMYNNKTRCTILADERSKPLGR	284							
QY	180	ANLAEVGMWTFEES-----RGVSEFTVPQMLVGFASFWMENPSTYMLCDOTS	230							
Db	285	ADLIATGEFTYFEKKCFASPTNCRVPSFKRPQMLVGFAPVMEVNPSTYMLCDOTN	344							
QY	231	PPETGQNFVCKSVWYNYNE	250							
Db	345	PPETGDFVCKSVWYNYNE	364							
RESULT 2										
AC	08T3B7	PRELIMINARY;	PRT;	1069	AA.					
DT	01-JUN-2002	(TReMBLrel. 21, Created)								
DT	01-OCT-2002	(TReMBLrel. 21, Last sequence update)								
DT	01-OCT-2003	(TReMBLrel. 25, Last annotation update)								
DE	Hypothetical protein C34G6.6;									
GN	Name=C34G6.6; ORFNames=C34G6.6;									
OS	Caenorhabditis elegans.									
OC	Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;									
OC	Rhabditidae; Peloderinae; Caenorhabditis.									
OX	NCBI_TaxID=6239;									
RP	[1]									
RP	SEQUENCE FROM N.A.									
RC	STRAIN=Bristol N2;									
RC	MEDLINE=99069613; PubMed=9851916;									
RX	WormBase Consortium;									
RG	"Genome sequence of the nematode C. elegans: a platform for									
RT	investigating biology. The C. elegans Sequencing Consortium."									
RL	Science 282:2012-2018 (1998).									
RN	[2]									
RP	SEQUENCE FROM N.A.									
RC	STRAIN=Bristol N2;									
RA	Langston Y., Rohlfing T.;									
RT	"The sequence of C. elegans cosmid C34G6.6";									
RL	Submitted (APR-1997) to the EMBL/GenBank/DBJ databases.									
RN	[3]									
RP	SEQUENCE FROM N.A.									
RC	STRAIN=Bristol N2;									
RA	Waterston R.;									
RL	Submitted (DEC-2002) to the EMBL/GenBank/DBJ databases.									
RN	[4]									
RP	SEQUENCE FROM N.A.									
RC	STRAIN=Bristol N2;									
RG	WormBase Consortium;									
RL	Submitted (SEP-2004) to the EMBL/GenBank/DBJ databases.									
EMBL	U97407; AAM15558.1; -;									
DR	WormBase; WBGene00016422; C34G6.6.									
DR	WormPep; C34G6.6b; CE30432.									
DR	InterPro; IPR001507; EndoglIn/CD105.									
DR	InterPro; IPR003014; PAN.									
DR	InterPro; IPR003609; Pan_app.									
DR	Pfam; PF00024; PAN; 6.									
DR	Pfam; PF00100; Zona_pellucida; 1.									
DR	SMART; SM00473; PAN_AP; 6.									
DR	SMART; SM00241; ZP; 1.									
DR	PROSITE; PS50948; PAN; 6.									
DR	Hypothetical protein.									

Query Match	Similarity	67.4%	Score 912.5	DB 2	Length 1069
Best Local	Similarity	66.2%	Pred. No. 8.6e-76		
Matches	172	Conservative	33	Mismatches	36
				Indels	19
				Gaps	5
QY	1	EOKILVGFAREVSADSVHRCLSACLNAFDTFFGECESVMYYPVDAECILNTEDELDRPD	60		
DB	114	EORILVGFARREVPAPANIQIMAAACLNAFDTFFGECESAMFYYPVQECILNTEDELDRPS	173		
QY	61	LFVNEHEPTVLYLNNNGCGCGCHNFPDNFKSGLINDDQFAIA-AQCAPVTOVAAEG	119		
DB	174	LFVESDDTVLYMDNCGAGPPL-VFNKY-----NQKTFPSQCYPPYITOTYIAVEG	225		
QY	120	ROLSDELHSEGEELSECBEELCTORLSVTANDFNCKSFMTSNLTRSCVLDSERSRPLGR	179		
DB	226	KQKNNEEDRLI-NVDLDSOCLACTORLSSINDFNCKSFMTNNKTRCTIILADESKPLGR	284		
QY	180	ANLAEVPGWYFES-----RGVPSFTVPQMLLVGFAPSPMENVPSYTMCLDOCTS	230		
DB	285	ADLITATBGFYFEKCKCFASPNTCHNVPSFKRVPQMLLVGFAPSPMENVPSYTMCLDOCTN	344		
QY	231	PPETGQNFVCKSVMYYYNE 250			
DB	345	PPETGDFGVCKSVMYYYNE 364			
RESULT 3					
ID	Q20167	PRELIMINARY;	PRT;	751 AA.	
AC	Q20167				
DT	01-NOV-1996	(TREMBLrel. 01, Created)			
DT	01-NOV-1996	(TREMBLrel. 01, Last sequence update)			
DT	01-JUN-2003	(TREMBLrel. 24, Last annotation update)			
DE		Hypothetical protein F38B11.4.			
CN		ORNames=F38B11.4;			
OS		Caenorhabditis elegans.			
OC		Eukaryota; Metazoa; Nematoda; Chromodorea; Rhabditida; Rhabditoidea;			
OC		Rhabditidae; Peloderinae; Caenorhabditis.			
OK	NCHI_TaxID=6239;				
RN	[1]				
RP	SEQUENCE FROM N.A.				
RC	STRAIN=Bristol N2;				
RX	MEDLINE=99069613; PubMed=9851916;				
RA	none;				
RT	"Genome sequence of the nematode C. elegans: A platform for				
RT	investigating biology."				
RL	Science 282:2012-2018 (1998).				
RL	[2]				
RP	SEQUENCE FROM N.A.				
RC	STRAIN=Bristol N2;				
RA	Matches P.;				
RL	Submitted (JAN-1996) to the EMBL/GenBank/DBJ databases.				
DR	EMBL; Z68342; CAA92773.1; -.				
DR	PIR; T21967; T21967.				
DR	WormBase; WBGene00009541; F38B11.4.				
DR	WormPep; F38B11.4; CE03296.				
DR	InterPro; IPR001507; Exodglin/CD105.				
DR	InterPro; IPR003014; PAN.				
DR	InterPro; IPR003609; Pan_app.				
DR	Pfam; PF00024; PAN; 4.				
DR	Pfam; PF00100; Zona pellucida; 1.				
DR	SMART; SM00473; PAN_AP; 4.				
DR	SMART; SM00241; ZP; 1.				
DR	PROSITE; PS50948; PAN; 4.				
KM	Hypothetical protein.				
SO	SEQUENCE 751 AA; 84391 MW; 92E4FE353CCBFN4 CRC64;				
Query Match 19.4%; Score 262; DB 2; Length 751;					
Best Local Similarity 27.4%; Pred. No. 1.7e-15;					
Matches 73; Conservative 37; Mismatches 98; Indels 58; Gaps 8;					
QY	2	OKILVGFAREVSADSVHRCLSACLNAFDTFFGECESVMYYPVDAECILNTEDELDRPD	61		

Db 131 OSILIGHAMKVLTVYDGLSDCLSKALSORSYDPLCKSAIYYETGECIMRDNKFIYPKL 190  
 Qy 62 FVDEHEDPVI-YLNNCGAGCEC-----HHFDFNFKTSGILNDOQFALAOCYAPYTOY 114  
 Db 191 FKTNILDLVDFYFNKNCADVSCKEPETHW----- 220  
 Qy 115 VAVGRGRLSDLDHSEFGLSECEBELCTQRLSTVANDFNCKSPMYSNLTRSCVLSDBS 174  
 Db 221 VRTEBYLIDSKDVIYESSDAOECNOLC-QNNKIGENFPCKAPATYSKQECHLTAESS 279  
 Qy 175 ----RPLGRANLAEPVGMWTFESRGVP-----SFTRVPMQLLVGFASFVMEVPSVT 222  
 Db 280 YVGHKGDRFNLAPLNFSEGEYFEKYLPTNLQICIBASFELVANRMT--SAVKTISALSOH 337  
 Qy 223 MCLDOCTSPPEPTGQNFVCKSVMYTY 248  
 Db 338 ECLSQCMNDGAR-----CSSATYFY 357

RESULT 4  
 ID 09XUES PRELIMINARY; PRT; 741 AA.  
 AC 09XUES;  
 DT 01-NOV-1999 (TREMBLrel. 12, Created)  
 DT 01-NOV-1999 (TREMBLrel. 12, Last sequence update)  
 DT 01-JUN-2003 (TREMBLrel. 24, Last annotation update)  
 DE Hypothetical protein F52B11.3.  
 GN ORFNames=F52B11.3;  
 OS Caenorhabditis elegans.  
 OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditioidea;  
 OC Rhabditidae; Pelodertinae; Caenorhabditis.  
 NCBI\_TaxID=6239;  
 RN [1]  
 RC SEQUENCE FROM N.A.  
 RC STRAIN=Bristol N2;  
 RX MEDLINE=99069613; PubMed=9851916;  
 RA none;  
 RT "Genome sequence of the nematode C. elegans: A platform for  
 RT investigating biology."  
 RL Science 282:2012-2018 (1998).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RP STRAIN=Bristol N2;  
 RC Matthews L.;  
 RL Submitted (NOV-1996) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; T22486; CAB05199.1; -.  
 DR PIR; T22486; T22486.  
 DR WormBase; WBGene0000926; F52B11.3.  
 DR WormPeP; F52B11.3; CE18723.  
 DR InterPro; IPR001507; Endoglin/CD105.  
 DR InterPro; IPR003014; PAN.  
 DR Pfam; PF00024; PAN; 4.  
 DR Pfam; PF00100; zona pellucida; 1.  
 DR SMART; SM00473; PAN\_AP; 4.  
 DR SMART; SM00241; ZP; 1.  
 DR PROSITE; PS50948; PAN; 4.  
 DR Hypothetical protein.  
 KW SEQUENCE 741 AA; 81738 MW; F6BC96FED7D8C15 CRC64;  
 SQ

Query Match 12.4%; Score 168; DB 2; Length 741;  
 Best Local Similarity 22.5%; Pred. No. 9.3e-07;  
 Matches 76; Conservative 38; Mismatches 100; Indels 124; Gaps 15;

Qy 5 LVGFARVVSADSVARCLASCLNADPTGFECEGVYVVDACLTINTEDRLRDLPLVD 64  
 Db 137 LEGFVKSVTVENRHCISACIKEYE---FVCKSVNHNHTSLCELSYEDKSKXP----- 188  
 Qy 65 EH---EDTVYIYDNNCAG-----C--EC 82  
 Db 169 THVRMSKIDYDNCILSRONRCPSGNLVKTTNFEIRYDHTQSVQAEQSYCLQKC 248  
 Qy 83 HHNFDNFKTSGILN-----DQG-----FALDAQCYAPYVTOY 114

Db 249 LDSLNTFCRSVFEVPEKNCIVSDEDTFSRADQGOVGVKDYEPICVAADLSSTCRQ 308  
 Qy 115 VAVE---GRQLSDLDHSEFGLSECEBELCTQRLSTVANDFNCKSPMYSNLTRSCVLSDBS 177  
 Db 309 AAFERFSGSIEGVVVAAGQVTTISDCLCFQNL-----NCKSINYDRTASSCFI-- 359  
 Qy 172 ERSRPLGR--ANLAEPVGMWTFESRGVPSF-----TRVPMQLLVGFA-- 211  
 Db 360 ---YAVGRQDHNITANDSMIDYEFNCESQFGMALCTNEGIRFTVNTKEPTGALYAAR 416  
 Qy 212 ----SFVMEVPSVTMCLDOCTSPPEPTGQNFVCKSVY 245  
 Db 417 FSTCSQVVENAKQISI-----TPPPTVYSSD--CGTVI 447

RESULT 5  
 ID 017347 PRELIMINARY; PRT; 522 AA.  
 AC 017347;  
 DT 01-JAN-1998 (TREMBLrel. 05, Created)  
 DT 01-JAN-1998 (TREMBLrel. 05, Last sequence update)  
 DT 01-OCT-2003 (TREMBLrel. 25, Last annotation update)  
 DE Hypothetical protein H42K12.3.  
 GN Name=H42K12.3; ORFNames=H42K12.3;  
 OS Caenorhabditis elegans.  
 OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditioidea;  
 OC Rhabditidae; Pelodertinae; Caenorhabditis.  
 NCBI\_TaxID=6239;  
 RN [1]  
 RC SEQUENCE FROM N.A.  
 RC STRAIN=Bristol N2;  
 RX MEDLINE=99069613; PubMed=9851916;  
 RG WormBase Consortium;  
 RT "Genome sequence of the nematode C. elegans: a platform for  
 RT investigating biology. The C. elegans Sequencing Consortium.";  
 RL Science 282:2012-2018 (1998).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RP STRAIN=Bristol N2;  
 RC Maggi L., Harper M.;  
 RA "The sequence of C. elegans fosmid H42K12."  
 RL Submitted (OCT-1997) to the EMBL/GenBank/DBJ databases.  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 RP STRAIN=Bristol N2;  
 RC Watson R.;  
 RL Submitted (NOV-2002) to the EMBL/GenBank/DBJ databases.  
 RN [4]  
 RP SEQUENCE FROM N.A.  
 RP STRAIN=Bristol N2;  
 RC Wilson R.;  
 RL Submitted (JUL-2003) to the EMBL/GenBank/DBJ databases.  
 RN [5]  
 RP SEQUENCE FROM N.A.  
 RP STRAIN=Bristol N2;  
 RG WormBase Consortium;  
 RL Submitted (SEP-2004) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AF026207; AAB71266.1; -.  
 DR PIR; T32444; T32444.  
 DR WormBase; WBGene00019272; H42K12.3.  
 DR WormPeP; H42K12.3; CE11614.  
 DR InterPro; IPR003014; PAN.  
 DR InterPro; IPR003609; Pan\_APP.  
 DR Pfam; PF00024; PAN; 2.  
 DR Pfam; PF00084; TSP1.  
 DR SMART; SM00473; PAN\_AP; 2.  
 DR SMART; SM00209; TSP1; 1.  
 DR PROSITE; PS50948; PAN; 2.  
 DR PROSITE; PS50092; TSP1; 1.  
 DR Hypothetical protein.  
 KW SEQUENCE 522 AA; 58738 MW; C4E804F9A0059180 CRC64;  
 SQ



DT 01-MAR-2004 (TReMBLrel). 26, Laet annotation update)

DE CG13207-PC (CG13207-pd).

GN Name=nompA; ORFNames=CG13207;

OS *Drosophila melanogaster* (fruit fly).

OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;  
OC Neoptera; Endopterygota; Diptera; Brachyera; Muscomorpha;  
OC Ephydroidea; Drosophilidae; Drosophila.

OX NCBI\_TaxID=7227;

XX

RN RN

RP SEQUENCE FROM N.A.

RX MEDLINE=20196006; PubMed=10731132; DOI=10.1126/science.287.5461.2185;  
RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,  
RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,  
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,  
RA Sutton G.G., Wortman J.E., Yandell M.D., Zhang Q., Chen L.X.,  
RA Brannon R.C., Rogers Y.H., Blazer R.G., Champe M., Pfeiffer B.D.,  
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Gabor G.L.,  
RA Arroll J.P., Agapayni A., An H.J., Andrews-Pfannkoch C., Baldwin D.,  
RA Ballew R.M., Basu A., Bayandale J., Bayraktiroglu L., Beasley E.M.,  
RA Beeson K.Y., Benos P.V., Bertman B.P., Bhandari D., Bolshakov S.,  
RA Borokova D., Botchan M.R., Bouck J., Brockstein P., Brotlier P.,  
RA Butts K.C., Busam D.A., Butler H., Criksten E., Center A., Chandra I.,  
RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,  
RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,  
RA Dodson K., Doup L.E., Downe M., Dugan-Rocha S., Dunkov B.C., Dunn P.,  
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferrieria S., Fleischman W.,  
RA Fodor C., Gabrielian A.E., Garg N.S., Gilbert W.M., Glasser K.,  
RA Glodet A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,  
RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,  
RA Hostali M., Houston K.A., Howland T.J., Wei M.H., Idegam C.A.,  
RA Uhalil M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,  
RA Kimmel B.E., Koditz C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,  
RA Laeko P., Lai Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,  
RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,  
RA Merkulov G., Mlshina N.V., Mobery C., Morris J., Moshrefi A.,  
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,  
RA Nelson D.R., Nelson K.A., Nixon K., Nusserm D.R., Pachet J.M.,  
RA Palazzolo M., Peltman G.S., Pan S., Pollard J., Puti V., Reese M.G.,  
RA Rainett K., Remington K., Saunders R.D., Scheefel F., Shen H.,  
RA Shue B.C., Siden-Klamos I., Simpson M., Skupski M.P., Smith T.,  
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,  
RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,  
RA Wang Z.Y., Wasserman D.A., Weinstein G.M., Weissbach J.,  
RA Williams S.M., Woodager, Worley K.C., Wu D., Yang S., Yao Q.A., Ye J.,  
RA Yeh R.F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,  
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,  
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;  
RT "The genome sequence of *Drosophila melanogaster*.";  
RL Science 287:2185-2195(2000).

RN RN

RP SEQUENCE FROM N.A.

RX MEDLINE=22426065; PubMed=12537572;  
RA Celniker S., Adams M., Champe M., Dugan S.P., Frise E., Hodgson A.,  
RA Patel S., Patil S., Hoskins R.A., Laverly T., Muzny D.M., Nelson C.R.,  
RA George R.A., Park S., Pfeiffer B.D., Richards S., Sodergren E.J.,  
RA Svirskas R., Tabor P.E., Wan K., Stapleton M., Sutton G.G., Venter C.,  
RA Weinstein G., Scherer S.E., Myers E.W., Gibbs R.A., Rubin G.M.;  
RT "Finishing a whole-genome shotgun: Release 3 of the *Drosophila*  
RT *melanogaster* euchromatic genome sequence.";  
RL Genome Biol. 3:RESEARCH0079-RESEARCH0079(2002).

RN RN

RP SEQUENCE FROM N.A.

RX MEDLINE=22426070; PubMed=12537573;  
RA Kaminker J.S., Bergman C.M., Kronmiller B., Carlson J., Svirskas R.,  
RA Patel S., Frise E., Wheeler D.A., Lewis S.E., Rubin G.M.,  
RA Ashburner M., Celniker S.E.;  
RT "The transposable elements of the *Drosophila melanogaster* euchromatin  
RT a genomics perspective.";  
RL Genome Biol. 3:RESEARCH0084-RESEARCH0084(2002).

RN RN

RP SEQUENCE FROM N.A.

RX MEDLINE=22426069; PubMed=12537572;  
RA

Accession	Score	Length	Matches	Mismatches	Indels	Gaps
RA Maira S., Crosby M.A., Muninger C.J., Matthews B.B., Campbell K.S., Raedecker P., Huang Y., Kinkner J.S., Millburn G.H., Prochnik S.E., Smith C.D., Tupy U.L., Whitfield E.J., Bayraktaroglu I., Berman B.P., Ra Battencourt B.R., Celniker S.E., de Grey A.D., Drysdale R.A., Harris N.L., Richter J., Russo S., Schroeder A.J., Shu S.Q., Stepieton M., Yamada C., Ashburner M., Gelbart W.M., Rubin G.M., Lewis S.E.;	10.24	1549	58	100	72	13
RT "annotation of the Drosophila melanogaster euchromatic genome: a RT systematic review.";	21.24	1549	43	100	72	13
RL Genome Biol. 3:RESEARCH0083-RESEARCH0083(2002).	21.24	1549	43	100	72	13
RP SEQUENCE FROM N.A.	21.24	1549	43	100	72	13
RG FlyBase;	21.24	1549	43	100	72	13
RL Submitted (SEP-2002) to the EMBL/GenBank/DBJ databases.	21.24	1549	43	100	72	13
RN [6]	21.24	1549	43	100	72	13
RP SEQUENCE FROM N.A.	21.24	1549	43	100	72	13
RG FlyBase;	21.24	1549	43	100	72	13
RL Submitted (MAR-2004) to the EMBL/GenBank/DBJ databases.	21.24	1549	43	100	72	13
DR EMBL; AEO03826; AAM68728.2; -	21.24	1549	43	100	72	13
DR FlyBase; FBGN016047; nompA.	21.24	1549	43	100	72	13
DR GO; GO:0016358; P:dendrite morphogenesis; TAS.	21.24	1549	43	100	72	13
DR GO; GO:0009592; P:detection of sound; IMP.	21.24	1549	43	100	72	13
DR GO; GO:0007605; P:perception of sound; IMP.	21.24	1549	43	100	72	13
DR InterPro; IPR001507; Endoglin/CD105.	21.24	1549	43	100	72	13
DR InterPro; IPR003014; PAN.	21.24	1549	43	100	72	13
DR InterPro; IPR003609; Pan_app.	21.24	1549	43	100	72	13
DR Pfam; PF00024; PAN; 4.	21.24	1549	43	100	72	13
DR Pfam; PF00100; Zona pellucida; 1.	21.24	1549	43	100	72	13
DR SMART; SM00473; PAN_AP; 4.	21.24	1549	43	100	72	13
DR SMART; SM00241; ZP; 1.	21.24	1549	43	100	72	13
DR PROSITE; PSS0948; PAN; 5.	21.24	1549	43	100	72	13
SJ SEQUENCE 1549 AA; 174236 MW; 5C65F4D8A2E7CD2C CRG64;	21.24	1549	43	100	72	13
Query Match	10.24	Score 138; DB 2; Length 1549;				
Best Local Similarity	21.24	Pred. No. 0.0014;				
Matches	58;	Conservative 43; Mismatches 100; Indels 72; Gaps 13				
QY 3 KILVGFARREVVSADSVHRCISACINAFDTGFCFECESVMYVYVDAECILNTEDELDLPDLF 62						
DB 164 KLPISDIKETIAANR-SDCEBCKLN---EFSFVCRSANFDSMTSCITLSRFTRRYHPIIM 219						
QY 63 VDEHEDIVYLDNNCAGECHMHDFNFTSGILNDQCAIAACVAPVTOVVAEGRQL 122						
DB 220 EDDNSD-VLENTCLNAB-----RRCGG-LAVFVKEENKRL 253						
QY 123 SD--ELDHSFEGELSECCELTQRLSYTANDFNCKSPMYGNLFRSCLSDERS-RPLGR 179						
DB 254 GGFPEVD-IFNNMTLEBCCQTC-----LRACKYCRSVFPPDQSCQCLTSBEDSISQKDD 307						
QY 180 ANIAAEVGMVTFE-----SRGVPSFTVPQMLVNG-FASF 213						
DB 308 ISISSPTHHFYDVLCLDNQRANDYPDWSVYTHLPSGGRBDPTAFORYNSRLGGEFHSF 367						
QY 214 VMENVSVYTMCLDOCTSPRPETGQNFVCKSMY 246						
DB 368 IYGR--SLSECLDECLRQ-----TSFOCRASY 393						
RESULT 9						
Q9BMD4 PRELIMINARY; PRT; 1549 AA.						
AC Q9BMD4;						
DT 01-JUN-2001 (TREMBLrel. 17, Created)						
DT 01-JUN-2001 (TREMBLrel. 17, Last sequence update)						
DT 01-MAR-2004 (TREMBLrel. 26, Last annotation update)						
DE No-mechanoreceptor potential A short isoform precursor.						
GN Name=nompA;						
OS Drosophila melanogaster (Fruit fly).						
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;						
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;						
OC Ephydroidea; Drosophilidae; Drosophila.						
NCBI_TaxID=7227;						
LN [1]						

```

RP SEQUENCE FROM N.A.
RX MEDLINE=21134724; PubMed=11239432; DOI=10.1016/S0896-6273(01)00215-X;
RA Chung Y.D., Zhu J., Han Y., Kernan M.J.;
RT "nompA encodes a PMS-specific, ZP domain protein required to connect
RT mechanosensory dendrites to sensory structures.";
RL Neuron 23:415-428(2001).
DR EMBL: AF34032; AAK09434.1; -.
DR FlyBase: FBgn0016047; nompA.
DR GO: GO:0016358; P.dendrite morphogenesis; TAS.
DR GO: GO:0009592; P.detection of sound; IMP.
DR GO: GO:0007605; P.perception of sound; IMP.
DR InterPro: IPR001507; Endoglin/CD105.
DR InterPro: IPR003014; PAN.
DR InterPro: IPR003609; Pan_app.
DR Pfam: PF00024; PAN; 4.
DR Pfam: PF00100; Zon pellucida; 1.
DR SMART: SM00473; PAN_AP; 4.
DR SMART: SM00241; ZP; 1.
DR PROSITE: PS0948; PAN; 5.
DR Receptor; Signal.
KW SIGNAL.
FT SIGNAL.
SQ SEQUENCE 1549 AA; 174270 MW; 78C31BACA39D5B93 CRC64;

Query Match 10.2%; Score 138; DB 2; Length 1549;
Best Local Similarity 21.2%; Pred. No. 0.0014;
Matches 58; Conservative 43; Mismatches 100; Indels 72; Gaps 13;

QY 3 KLVGFARVVSADSVHRCISACLNAPDFPFGECSVMYYPVDAECILNTEBRLDRPLF 62
DB 164 KLPISDIKEITANR-SDCEBKCLN--EFSFVCRSANFDSTMRCTLSRFTRRHPILM 219
QY 63 VDEHEDTVIYLDNNCAGCECHMHFDFNFKTSGILNDQFALAAQCVAPYVTVAVVAGRQL 122
DB 220 EDDPNSD--YLENTCLNAB-----RRCDG--LAVFVKEENKRL 253
QY 123 SD--ELDHSFEGLESECELCCTQRLSYTANDFNCKSFMYNLTSSCVLSDBRS-RPLGR 179
DB 254 GGPFEVD-IFNMNMLEECQTM-----LRAEKYFRSVFDDQSKQCLISEBDSISQKDD 307
QY 180 ANLAEPVPGWTFPE-----SRGVPSFTRVPQMLLVG-FAF 213
DB 308 ISISSPFTHFYDVLCLNDQANDYPDNTSVTSHLSSGRRPDTAFQRYNSRLGGFFHSE 367
QY 214 VMENVPSVTMCLDOCTSPPEPTGONFVCKSVMY 246
DB 368 ITGR--SLSECLDECLRO-----TSFQCRSAVY 393
QY 368 ITGR--SLSECLDECLRO-----TSFQCRSAVY 393

RESULT 10
Q9BMD5 PRELIMINARY; PRT; 1557 AA.
AC Q9BMD5;
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DE 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE No-mechanoreceptor potential A long isoform precursor.
GN Name=nompA;
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OC NCBI_TaxId=7227;
RN NCBI_TaxId=7227;
RX MEDLINE=21134724; PubMed=11239432; DOI=10.1016/S0896-6273(01)00215-X;
RA Chung Y.D., Zhu J., Han Y., Kernan M.J.;
RT "nompA encodes a PMS-specific, ZP domain protein required to connect
RT mechanosensory dendrites to sensory structures.";
RL Neuron 23:415-428(2001).
DR EMBL: AF34032; AAK09434.1; -.
DR FlyBase: FBgn0016047; nompA.
DR GO: GO:0016358; P.dendrite morphogenesis; TAS.
DR GO: GO:0009592; P.detection of sound; IMP.
DR GO: GO:0007605; P.perception of sound; IMP.
DR Receptor; Signal.
KW SIGNAL.
FT SIGNAL.
SQ SEQUENCE 1557 AA; 175290 MW; 134BBBA89245183 CRC64;

Query Match 10.2%; Score 138; DB 2; Length 1557;
Best Local Similarity 21.2%; Pred. No. 0.0014;
Matches 58; Conservative 43; Mismatches 100; Indels 72; Gaps 13;

QY 3 KLVGFARVVSADSVHRCISACLNAPDFPFGECSVMYYPVDAECILNTEBRLDRPLF 62
DB 164 KLPISDIKEITANR-SDCEBKCLN--EFSFVCRSANFDSTMRCTLSRFTRRHPILM 219
QY 63 VDEHEDTVIYLDNNCAGCECHMHFDFNFKTSGILNDQFALAAQCVAPYVTVAVVAGRQL 122
DB 220 EDDPNSD--YLENTCLNAB-----RRCDG--LAVFVKEENKRL 253
QY 123 SD--ELDHSFEGLESECELCCTQRLSYTANDFNCKSFMYNLTSSCVLSDBRS-RPLGR 179
DB 254 GGPFEVD-IFNMNMLEECQTM-----LRAEKYFRSVFDDQSKQCLISEBDSISQKDD 307
QY 180 ANLAEPVPGWTFPE-----SRGVPSFTRVPQMLLVG-FAF 213
DB 308 ISISSPFTHFYDVLCLNDQANDYPDNTSVTSHLSSGRRPDTAFQRYNSRLGGFFHSE 367
QY 214 VMENVPSVTMCLDOCTSPPEPTGONFVCKSVMY 246
DB 368 ITGR--SLSECLDECLRO-----TSFQCRSAVY 393
QY 368 ITGR--SLSECLDECLRO-----TSFQCRSAVY 393

RESULT 11
Q9V5X3 PRELIMINARY; PRT; 1557 AA.
AC Q9V5X3;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DE 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE CG13207-PA (CG13207-pb).
GN Name=nompA; ORFNames=CG13207;
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OC NCBI_TaxId=7227;
RN NCBI_TaxId=7227;
RX MEDLINE=20196006; PubMed=10731132; DOI=10.1126/science.287.5461.2185;
RA Adams M.D., Celinker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutcliffe G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA Brandon R.C., Rogers Y.H., Blazer R.G., Champe M., Pfeiffer B.D.,
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Gabor G.L.,
RA Abail J.F., Agbayani A., An H.J., Andrews-Pfannkoch C., Baldwin D.,
RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA Borokova D., Borchan M.R., Bouck J., Brokstein P., Brotlier P.,
RA Butris K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Doudon K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Durkin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
RA Foster C., Gabrieli A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA Glodok A., Gong F., Gottrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,

```

RA Hostin D., Houston K.A., Howland T.J., Wei M.H., Ibegwam C.,  
RA Jatali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Kechum K.A.,  
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,  
RA Liao P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,  
RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,  
RA Mekulov G., Mishina N.V., Modary C., Morris J., Moshrefi A.,  
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,  
RA Nelson S.R., Nelson K.A., Nixon K., Nuskern D.R., Pacleb J.M.,  
RA Palazzolo M., Pletman G.S., Pan S., Pollard J., Puri V., Reese M.G.,  
RA Reinert K., Remington K., Saunders R.D., Scheeler F., Shen H.,  
RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,  
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,  
RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,  
RA Wang Z.Y., Weissman D.A., Weinstein G.M., Weisenbach J.,  
RA Williams S.M., Woodgerf, Worley K.C., Wu D., Yang S., Yao Q.A., Ye J.,  
RA Yeh R.F., Zaveri J.S., Zhan W., Zhang G., Zhao Q., Zheng L.,  
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,  
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;  
RT "The genome sequence of *Drosophila melanogaster*.";  
RL Science 287:2185-2195 (2000).  
[2]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=22426065; Pubmed=12537568;  
RA Celniker S.E., Wheeler D.A., Krommiller B., Carlson J.W., Halpern A.,  
RA Patel S., Adams M., Champe M., Dugan S.P., Friese E., Hodgson A.,  
RA George R.A., Hoskins R.A., Laverly T., Muzny D.M., Nelson C.R.,  
RA Pacleb J.M., Park S., Pfeiffer B.D., Richards S., Sodergren E.J.,  
RA Svirskas R., Taber P.E., Wan K., Stapleton M., Sutton G.G., Venter C.,  
RA Weinstein G., Scherer S.E., Myers E.W., Gibbs R.A., Rubin G.M.;  
RT "Finishing a whole-genome shotgun: Release 3 of the *Drosophila*  
RT *melanogaster* euchromatic genome sequence.";  
RL Genome Biol. 3:RESEARCH0079-RESEARCH0079(2002).  
[3]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=22426070; Pubmed=12537573;  
RA Kaminker J.S., Bergman C.M., Krommiller B., Carlson J., Svirskas R.,  
RA Patel S., Friese E., Wheeler D.A., Lewis S.E., Rubin G.M.,  
RA Ashburner M., Celniker S.E.;  
RT "The transposable elements of the *Drosophila melanogaster* euchromatin:  
RT a genomic perspective.";  
RL Genome Biol. 3:RESEARCH0084-RESEARCH0084(2002).  
[4]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=22426069; Pubmed=12537572;  
RA Miara S., Crosby M.A., Mungall C.J., Matthews B.B., Campbell K.S.,  
RA Hrdceky P., Huang Y., Kaminker J.S., Millburn G.H., Prochuk S.E.,  
RA Smith C.D., Tupy J., Whitfield E.J., Bayraktaroglu U., Berman B.P.,  
RA Bertencourt B.R., Celniker S.E., de Grey A.D., Drysdale R.A.,  
RA Harris N.L., Richter J., Russo S., Schroeder A.J., Shu S.Q.,  
RA Stapleton M., Yamada C., Ashburner M., Gelbart W.M., Rubin G.M.,  
RA Lewis S.E.;  
RT "Annotation of the *Drosophila melanogaster* euchromatic genome: a  
RT systematic review.";  
RL Genome Biol. 3:RESEARCH0083-RESEARCH0083(2002).  
[5]  
RP SEQUENCE FROM N.A.  
RG FlyBase;  
RL Submitted (SEP-2002) to the EMBL/GenBank/DBJ databases.  
RN [6]  
RP SEQUENCE FROM N.A.  
RG FlyBase;  
RL Submitted (MAR-2004) to the EMBL/GenBank/DBJ databases.  
RN EMBL; AE003826; AAF58668.2; -  
DR FlyBase; FBgn0016047; nompA.  
DR GO; GO:0016358; P:dendrite morphogenesis; TAS.  
DR GO; GO:0009592; P:detection of sound; IMP.  
DR GO; GO:0007605; P:perception of sound; IMP.  
DR InterPro; IPR001507; Endoglin/CD105.  
DR InterPro; IPR003014; PAN.  
DR InterPro; IPR003609; Pan\_app.  
DR Pfam; PF00024; PAN; 4.  
DR Pfam; PF00100; Zon\_pellucida; 1.  
DR SMART; SM00473; PAN\_AP; 4.

DR SMART; SM00241; ZP; 1.  
DR PROSITE; PS50948; PAN; 5.  
SQ SEQUENCE 1557 AA; 175256 MW; E060906D48DB77DF CRC64;  
Query Match 10.2%; Score 138; DB 2; Length 1557;  
Best Local Similarity 21.2%; Pred. No. 0.0014;  
Matches 58; Conservative 43; Mismatches 100; Indels 72; Gaps 13;  
QY 3 KILVGFARVVSAQSVHRCASACINAFDTFGFEGESVWYVVDACILNTEDRLPDLF 62  
DB 164 KLPISDIKEITNAAR-SDCEKCLN--EFSFVCRSANFSDTMSCTLSRTFRTHPELM 219  
QY 63 VDEHEDVITYLDNNCAGCECHMFDPNFTSGILNDQQAACYPAYVTVYVAVEGROL 122  
DB 220 EDDNSD-YLENCLNAB-----RRCDG--LAFVKEENRL 253  
QY 123 SD--ELDHSFEGELSGEEELCTORLSYANDPNCCKSMYNLTRSCYLSDBRS-RLPLGR 179  
DB 254 GGPPEVD-IFNNMTLEECQTC-----LRAEKYCRSVFEPDQSKCILSEDSISQKDD 307  
QY 180 ANLAEVPGWTFE-----SRGVPSFTVRPQMLVVG-FASF 213  
DB 308 ISISSPTTHFFYDVLCDNQRANDYPDNTSHLFSGGRPDYTAFORYRNSRLGGSEFHS 367  
QY 214 VMENVPSVTMCLDOCTSPPEPTGONFVCKSVY 246  
DB 368 ITRG-SISECLDECLNQ-----TSFQCRSAVY 393  
RESULT 12  
ID Q21782 PRELIMINARY; PRT; 351 AA.  
AC Q21782;  
DT 01-NOV-1996 (TrEMBLrel. 01. Created)  
DT 01-NOV-1996 (TrEMBLrel. 01. Last sequence update)  
DT 01-MAR-2004 (TrEMBLrel. 26. Last annotation update)  
DE Hypothetical protein R07A4.4.  
GN ORFNames=R07A4.4;  
OS Caenorhabditis elegans.  
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditidae;  
OC Rhabditidae; Pelodierinae; Caenorhabditis.  
OX NCBI\_TaxID=6239;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=Bristol N2;  
RX MEDLINE=98063613; Pubmed=9851916;  
RA none;  
RT "Genome sequence of the nematode *C. elegans*: A platform for  
RT investigating biology.";  
RL Science 282:2012-2018 (1998).  
[2]  
RP SEQUENCE FROM N.A.  
RC STRAIN=Bristol N2;  
RA Cottage A.;  
RL Submitted (NOV-1995) to the EMBL/GenBank/DBJ databases.  
DR EMBL; Z67756; CA931764.1; -  
DR PIR; T23990; T23990.  
DR WormBase; WBGene00011075; R07A4.4.  
DR WormPeP; R07A4.4; CE03533.  
DR InterPro; IPR003014; PAN.  
DR InterPro; IPR003609; Pan\_app.  
DR InterPro; IPR008266; Tyr\_kinase\_AS.  
DR Pfam; PF00024; PAN; 2.  
DR SMART; SM00473; PAN\_AP; 2.  
DR PROSITE; PS50948; PAN; 2.  
DR PROSITE; PS00109; PROTEIN\_KINASE\_TYR; UNKNOWN\_1.  
KW Hypothetical protein.  
SQ SEQUENCE 351 AA; 39349 MW; C9E0BEC36612F80 CRC64;  
Query Match 9.3%; Score 126; DB 2; Length 351;  
Best Local Similarity 25.0%; Pred. No. 0.0029;  
Matches 45; Conservative 31; Mismatches 72; Indels 32; Gaps 7;



QY 95 LINDQO-----FA-----IAQCAPVYTOVAVAGEBQLSDELDHSPFEGLESCBETL 141  
 DB 8 INQOQOQFNDTALNFAVEUPLIPDPCPLGHSRQIIEGVEASREATITTVAILLEQCUA 67  
 QY 142 CTQRLSVTANDFN---CKSFVMSNLTBSCVLSDESRPLGRANLAEPVGMWTFESRGVP 197  
 DB 68 C-RVSTVADSGRLPLRGSAHFNRATRCQSVSDAINPNGVLEKPNQNIYIEKICIP 125  
 QY 198 S-----FTRVQMLLVGFASFVMEVNSVTMCLDOQCSPPPETQONFVCSVMY 247  
 DB 126 DTVLPWSCDVPRRIPOHILHGLASEVI-SVASENECVLECIX-AKTLRSVACHSILHY 182  
 QY RESULT 13  
 ID Q9V9X1 PRELIMINARY, PRT, 774 AA.  
 Q9V9X1  
 AC Q9V9X1;  
 DT 01-MAY-2000 (TrEMBLrel. 13, Created)  
 DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)  
 DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)  
 DE CG12063-PA.  
 GN ORFNames=CG12063;  
 OS Drosophila melanogaster (Fruit fly).  
 OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;  
 OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;  
 OC Ephydroidea; Drosophilidae; Drosophila.  
 NCBI\_TaxID=7227;  
 RX MEDLINE=20196006; PubMed=10731132; DOI=10.1126/science.287.5461.2185;  
 RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,  
 RA Amaratides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,  
 RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,  
 RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,  
 RA Brandon R.C., Rogers Y.H., Blaziel R.G., Champe W., Pfeiffer B.D.,  
 RA Wan K.H., Doyle C., Baxter A.G., Helt G., Nelson C.R., Gaber G.L.,  
 RA Abril J.F., Agbayani A., An H.J., Andrews-Pfannkoch C., Baldwin D.,  
 RA Ballew R.M., Basu P.A., Baxendale J., Bayraktaroglu L., Beasley E.M.,  
 RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolhaver S.,  
 RA Borokova D., Botchan M.R., Bouck J., Brokstein P., Brotlier P.,  
 RA Burris K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,  
 RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,  
 RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,  
 RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,  
 RA Durbin K.J., Evangelista C.C., Ferraz C., Ferriera S., Fleischmann W.,  
 RA Foster C., Gabriellian A.E., Garg N.S., Gelbart W.M., Glasner K.,  
 RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,  
 RA Harris N.L., Harvey D., Helman T.J., Hernandez J.R., Houck J.,  
 RA Hostin D., Houston K.A., Howland T.J., Wei M.H., Ibegwam C.,  
 RA Jajal M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,  
 RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,  
 RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,  
 RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,  
 RA Merkulov G., Mishina N.V., Mobarry C., Morris J., Mostrel A.,  
 RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,  
 RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,  
 RA Palazolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,  
 RA Reinert K., Remington K., Saunders R.D., Scheeler F., Shen H.,  
 RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,  
 RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,  
 RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,  
 RA Wang Z.Y., Wassarman D.A., Weinstein G.M., Weisenbach J.,  
 RA Williams S.M., Woodagst, Worley K.C., Wu D., Yang S., Yao Q.A., Ye J.,  
 RA Yeh R.F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,  
 RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,  
 RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;  
 RA "The genome sequence of Drosophila melanogaster.";  
 RA Science 287:2185-2195(2000).  
 RN  
 RN SEQUENCE FROM N.A.  
 RX MEDLINE=22426065; PubMed=12537568;  
 RA Celniker S.E., Wheeler D.A., Kronmiller B., Carlson J.W., Halpern A.,  
 RA Patel S., Adams M., Champe M., Dugan S.P., Frise E., Hodgson A.,

RA George R.A., Hoskine R.A., Lavery T., Muzny D.M., Nelson C.R.,  
 RA Pacleb J.M., Park S., Pfeiffer B.D., Richards S., Sodergren E.J.,  
 RA Svirskas R., Tabor P.E., Wan K., Stapleton M., Sutton G.G., Venter C.,  
 RA Weinstein R., Scherer S.E., Myers E.W., Gibbs R.A., Rubin G.M.,  
 RA "Finishing a whole-genome shotgun: Release 3 of the Drosophila  
 RT melanogaster euchromatic genome sequence.";  
 RL Genome Biol. 3:RESEARCH0079-RESEARCH0079(2002).  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 RA MEDLINE=22426070; PubMed=12537573;  
 RA Kaminker J.S., Bergman C.M., Kronmiller B., Carlson J., Svirskas R.,  
 RA Patel S., Frise E., Wheeler D.A., Lewis S.E., Rubin G.M.,  
 RA Ashburner M., Celniker S.E.;  
 RA "The transposable elements of the Drosophila melanogaster euchromatin:  
 RT a genomic perspective.";  
 RL Genome Biol. 3:RESEARCH0084-RESEARCH0084(2002).  
 RN [4]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=22426069; PubMed=12537572;  
 RA Misra S., Crosby M.A., Mungall C.J., Matthews B.B., Campbell K.S.,  
 RA Hradecky P., Huang Y., Kaminker J.S., Milburn G.H., Prochnik S.E.,  
 RA Smith C.D., Tupy J.L., Whitfield E.J., Bayraktaroglu L., Berman B.P.,  
 RA Betencourt B.R., Celniker S.E., de Grey A.D., Drysdale R.A.,  
 RA Harris N.L., Richter J., Ruso S., Schroeder A.J., Shu S.Q.,  
 RA Stapleton M., Yamada C., Ashburner M., Gelbart W.M., Rubin G.M.,  
 RA Lewis S.E.;  
 RA "Annotation of the Drosophila melanogaster euchromatic genome: a  
 RT systematic review.";  
 RL Genome Biol. 3:RESEARCH0083-RESEARCH0083(2002).  
 RN [5]  
 RP SEQUENCE FROM N.A.  
 RG FlyBase;  
 RL Submitted (SEP-2002) to the EMBL/GenBank/DBJ databases.  
 RN [6]  
 RP SEQUENCE FROM N.A.  
 RG FlyBase;  
 RL Submitted (MAR-2004) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AB003778; AAF51158.3; -  
 DR FlyBase; FBgn0039851; CG12063.  
 DR InterPro; IPR001507; Endoglin/CD105.  
 DR InterPro; IPR003014; PAN.  
 DR InterPro; IPR003609; Pan\_app.  
 DR Pfam; PF00024; PAN; 2.  
 DR Pfam; PF00100; Zona\_pellucida; 1.  
 DR SMART; SM00473; PAN\_AP; 3.  
 DR SMART; SM00241; ZP; 1.  
 DR PROSITE; PS50948; PAN; 3.  
 SQ SEQUENCE 774 AA; 83881 MW; 997A3B85350CECCC CRC64;  
 Query Match 9.2%; Score 125; DB 2; Length 774;  
 Best local similarity 25.6%; Pred. No. 0.0098;  
 Matches 62; Conservative 25; Mismatches 93; Indels 62; Gaps 13;  
 QY 21 CTSACLAADFTRGFCESVMTYVDAECL-NTEDRLDRPLFVDEHEDVIYLDNNKA 78  
 DB 173 CLEAC-----QANESCSAVNY--ETGLCVMEFSTYDOLP-GSLRSQPVFTVYAKSCF 224  
 QY 79 G---CECHMHFQNFRTSGIINDQCPAIAAGCAPVYTOVAVAGEBQLSDELDHSPFEGLE 135  
 DB 225 GVRPCKAKWICDR-----VQGRLEPRKAKSGSVATR 256  
 QY 136 SECEELCTQRLSVTANDFNCKSFVMSNLTBSCVLSDESRPLGRANLAEPVGMWTFESR 194  
 DB 257 RPCIIELC-----LGEITFCREANYVASGCLSDMDRITLSDBANIAAYGADYLENN 311  
 QY 195 GV--PS-----FTRVQMLLVGFASFVMEVNSVTMCLDOQCSPPPETQONFVCSVMY 248  
 DB 312 CAEBSKLCFEXRAGRIKTVDS-VHONVOTLDECRDLCTAP-----FRCHS--YDY 362  
 QY 249 NE 250  
 DB 363 NE 364



```

RESULT 14
O94248 PRELIMINARY; PRT: 1127 AA.
ID O94248
AC O94248;
DT 01-FEB-1997 (TREMBLrel. 02, Created)
DT 01-FEB-1997 (TREMBLrel. 02, Last sequence update)
DT 01-JUN-2003 (TREMBLrel. 24, Last annotation update)
DE Hypothetical protein H03E18.1.
GN Name=H03E18.1; ORFName=H03E18.1;
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoa; Chromadorea; Rhaditida; Rhaditoidae;
OC Rhaditidae; Peloderinae; Caenorhabditis.
NCBI_TaxID=6239;
RX MEDLINE=20196006; PubMed=9851916;
SEQUENCE FROM N.A.
RA STRAIN=Bristol N2;
RA MEDLINE=99069613; PubMed=9851916;
RA WormBase Consortium;
RT "Genome sequence of the nematode C. elegans: a platform for
investigating biology. The C. elegans Sequencing Consortium."
RL Sequence 282:2012-2018(1998).
RA Bentley D.;
RA "The sequence of C. elegans fosmid H03E18."
RL Submitted (SEP-1996) to the EMBL/GenBank/DBJ databases.
RA [3]
SEQUENCE FROM N.A.
RA STRAIN=Bristol N2;
RA WormBase Consortium;
RA Waterston R.;
RA Submitted (NOV-2002) to the EMBL/GenBank/DBJ databases.
RA [4]
SEQUENCE FROM N.A.
RA STRAIN=Bristol N2;
RA WormBase Consortium;
RL Submitted (SEP-2004) to the EMBL/GenBank/DBJ databases.
DR EMBL; U67947; AB07557.1; -.
DR PIR; T25804; T25804.
DR WormBase; WBGene00019148; H03E18.1.
DR WormPep; H03E18.1; CE11558.
DR InterPro; IPR003014; PAN.
DR InterPro; IPR003609; Pan_app.
DR Pfam; PF00024; PAN; 5.
DR SMART; SM00473; PAN AP; 5.
DR PROSITE; PS00948; PAN; 4.
DR Hypothetical protein.
KW SEQUENCE 1127 AA; 124650 MW; E7A7608E10EAFPP2 CRC64;
Query Match 9.2%; Score 124.5; DB 2; Length 1127;
blast Local Similarity 22.3%; Pred. No. 0.017;
Matches 44; Conservative 37; Mismatches 75; Indels 41; Gaps 9;
OY 1 BCKIVGFARVNSADSVARCASACLNAPDT--FGEEGSVWYVVDACLTNTEDRDR 58
DB 433 DDLHNVASVAGGEHMSIEECQCMANKSKTSRHYECSATYHAERCCILNLEDRNIK 492
OY 59 PDLFVDEHD-TVIVLDNNC-----AGCECHWHP-----NFKTSGIL 95
DB 433 SLLFPEKQPTDVVVYITGLACEVDETVKSGLALRECRKVVATGEPFKETSTKNGIK 552
OY 96 NDQQAIAACYAPYVTVYVAVEGRQSDDELHDSFEGLESCCELCTQRLSVTANDPNC 155
DB 553 SD-----DCYVE-LINPFLV-LEGTALIVE-----TAVTPECKCKCAEGHKLVEE--C 596
OY 156 KSEFVSNLTRSCVLSDE 172
DB 597 ASFLYYDSKTKLINQ 613
RESULT 15
O9VAG2

```

```

ID O9VAG2 PRELIMINARY; PRT: 744 AA.
AC O9VAG2;
DT 01-MAY-2000 (TREMBLrel. 13, Created)
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
DT 25-OCT-2004 (TREMBLrel. 28, Last annotation update)
DE CG7802-PA (CG7802-pb) (GH22837p).
GN ORFName=CG7802;
OS Drosophila melanogaster (Fruit fly).
OC Neoptera; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Ephydroidea; Ephydroptera; Diptera; Brachycera; Muscomorpha;
OC NCBI_TaxID=7227;
RX MEDLINE=20196006; PubMed=10731133; DOI=10.1126/science.287.5461.2185;
SEQUENCE FROM N.A.
RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Amanatides P.G., Scherer S.E., Li P.W., Hoeklin R.A., Galie R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA Brandon R.C., Rogers Y.H., Blazer V.G., Champe M., Pfeiffer B.D.,
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Gabot G.L.,
RA Abrita J.F., Agbayani A., An H.J., Andrews-Pfannkoch C., Baldwin D.,
RA Ballew R.M., Baau A., Baxendale J., Bayraktaroglu L., Beasley B.M.,
RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brotlier P.,
RA Burtis K.C., Busam D.A., Butler H., Cadiet L., Center A., Chandra I.,
RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA de Pablo B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dudson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
RA Foeller C., Gabriellian A.E., Garg N.S., Gelbart W.M., Glaeser K.,
RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
RA Hoscut D., Houston K.A., Howland T.O., Wei M.H., Ibegyan C.,
RA Jatali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA Laeko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Nusser D.R., Pacle J.M.,
RA Palazolo M., Peltman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA Reinert K., Remington K., Saunders R.D., Scheeler F., Shen H.,
RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA Spier B., Spradling A.C., Stapleton M., Strong R., Sun B.,
RA Svirskaes R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA Wang Z.Y., Wassarman D.A., Weisscock G.M., Weissbach J.,
RA Williams S.M., Woodgerter, Worley K.C., Wu D., Yang S., Yao Q.A., Ye J.,
RA Yen R.F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RA "The genome sequence of Drosophila melanogaster."
RL Sequence 287:2185-2195(2000).
[2]
SEQUENCE FROM N.A.
RA MEDLINE=22426065; PubMed=12537568;
RA Celniker S.E., Wheeler D.A., Krommiller B., Carlson J.W., Halpern A.,
RA Patel S., Adams M., Champe M., Dugan S.P., Frise E., Nelson A.,
RA George R.A., Hoeklin R.A., Laverly T., Muzny D.M., Nelson C.R.,
RA Pacle J.M., Park S., Pfeiffer B.D., Richards S., Sodergren E.J.,
RA Svirskaes R., Taber P.E., Wan K., Stapleton M., Sutton G.G., Venter C.,
RA Weinstein G., Scherer S.E., Myers E.W., Gibbs R.A., Rubin G.M.;
RA "Finishing a whole-genome shotgun Release 3 of the Drosophila
melanogaster euchromatic genome sequence."
RL Genome Biol. 3:RESEARCH0079-RESEARCH0079(2002).
[3]
SEQUENCE FROM N.A.
RA MEDLINE=22426070; PubMed=12537573;
RA Kaminker J.S., Bergman C.M., Krommiller B., Carlson J., Svirskaes R.,
RA Patel S., Frise E., Wheeler D.A., Lewis S.E., Rubin G.M.,
RA Ashburner M., Celniker S.E.;
RA "The transposable elements of the Drosophila melanogaster euchromatin:
a genomics perspective."

```

Genome Biol. 3:RESEARCH0084-RESEARCH0084(2002).

14

SEQUENCE FROM N.A.

Medline=22426069; PubMed=12537572;

Miera S., Crosby M.A., Mungall C.J., Matthews B.B., Campbell K.S.,

Hirdeczy P., Huang Y., Kaminker J.S., Milburn G.H., Prochuk S.E.,

Smith C.D., Tupy J.L., Whitfield E.J., Bayraktaroglu L., Berman B.P.,

Bettencourt B.R., Celinker S.E., de Grey A.D., Drysdale R.A.,

Harris N.L., Richter J., Russo S., Schroeder A.J., Shu S.Q.,

Stapleton M., Yamada C., Ashburner M., Gelbart W.M., Rubin G.M.,

Lewis S.E.;

"Annotation of the Drosophila melanogaster euchromatic genome: a

systematic review."

Genome Biol. 3:RESEARCH0083-RESEARCH0083(2002).

15

SEQUENCE FROM N.A.

FlyBase;

Submitted (SEP-2002) to the EMBL/GenBank/DBJ databases.

16

SEQUENCE FROM N.A.

Submitted (MAR-2004) to the EMBL/GenBank/DBJ databases.

17

SEQUENCE FROM N.A.

Stapleton M., Brockstein P., Hong L., Agbayani A., Carlson J.,

Champe M., Chavez C., Dorsett V., Dresnek D., Farfan D., Frise E.,

George R., Gonzalez M., Guatin H., Krommiller B., Li P., Liao G.,

Miranda A., Mungall C.J., Nunoo J., Pacleb J., Paragae V., Park S.,

Patel S., Phuanavong S., Wan K., Yu C., Lewis S.E., Rubin G.M.,

Celinker S.;

Submitted (APR-2002) to the EMBL/GenBank/DBJ databases.

18

EMBL; AE003771; AA56948.1; -

EMBL; AY04725; AAM1078.1; -

19

FlyBase; FBgn0039704; CG7802.

20

InterPro; IPR001507; Endoglin/CD105.

21

InterPro; IPR003014; PAN.

22

InterPro; IPR003609; Pan\_app.

23

Pfam; PF00024; PAN; 3.

24

Pfam; PF00100; Zona\_dellucida; 1.

25

SMART; SM00473; PAN\_AP; 3.

26

SMART; SM00241; ZP; 1.

27

ProSite; PS0948; PAN; 3.

28

SEQUENCE 744 AA; 83086 MW; FD2403A63F5FFBE CRC64;

Query Match 8.9%; Score 121; DB 2; Length 744;

Best Local Similarity 23.4%; Pred. No. 0.022;

Matches 64; Conservative 29; Mismatches 102; Indels 78; Gaps 14;

5 LVGEAFENVSA-----DS-----VHRCISACINAFDTFGFECESVWYYPVDAEC 48

117 LVGE--ELVIGYVSAPKELKMDSQGTMLTDCDCTCK-----NKTQSVNY--ETGLC 167

49 ILTNEBRLDRP-LFVDEHEDTVIYLNNNCAG--CECHWHEFNEKTSGLINDQFAIAA 104

168 VLFSAHADQLTGALTKSQFPVFTIYAQKSCIAVKCSRAWYVDR----- 211

105 QCVAPVYTVQYVAVSGROLSDLEHDSFEBLELSECEELCTQRLSVYANDFNCKSPYSLTJ 164

212 -----VQNYKLTKEVKTIVSVASRRECELC-----IGENDFTCRSANYDRTS 254

165 RSCVLSDESRHPLGRANLAEV-PGWTYPESRGVP-----SFTRPOMLLVGFASFVEN 217

255 GACELSELDRLLTLAGSQFQVNDGSDYLENHCVEBPNNKLCFKRPLRGLIKTVDS-VYQE 313

218 VPSYTNCLDOCTSPPEPTGQNFVCKSVWYYTNE 250

314 VSSIDECEELCLNSP-----YRCHS--YDND 338

Search completed: March 31, 2005, 02:06:06  
Job time : 180 secs